

Upper functions for positive random functionals. General setting and Gaussian random functions.

Oleg Lepski

*Laboratoire d'Analyse, Topologie, Probabilités
Université Aix-Marseille
39, rue F. Joliot-Curie
13453 Marseille, France
e-mail: lepski@cmi.univ-mrs.fr*

Abstract: In this paper we are interested in finding upper functions for a collection real-valued random variables $\{\Psi(\chi_\theta), \theta \in \Theta\}$. Here $\{\chi_\theta, \theta \in \Theta\}$ is a family of continuous random mappings, Ψ is a given sub-additive positive functional and Θ is a totally bounded subset of a metric space. We seek a non-random function $U : \Theta \rightarrow \mathbb{R}_+$ such that $\sup_{\theta \in \Theta} \{\Psi(\chi_\theta) - U(\theta)\}_+$ is "small" with prescribed probability. We apply the results obtained in the general setting to the variety of problems related to gaussian random functions and empirical processes.

AMS 2000 subject classifications: Primary 60E15; secondary 62G07, 62G08.

Keywords and phrases: upper function, empirical processes, gaussian random function, metric entropy, doubling measure.

1. Introduction

The main objective of this paper is to look from a novel point of view at some phenomena arising in different areas of probability theory and mathematical statistics. We will try to understand what is common between classical probabilistic results, such as the law of iterated logarithm for example, and well-known problem in adaptive estimation called price to pay for adaptation. Why do two different kinds of this price exist? What relates exponential inequalities for M -estimators, so-called uniform-in-bandwidth consistency in density or regression model and the bounds for modulus of continuity of gaussian random functions defined on a metric space equipped with doubling measure?

It turned out that all these and many other problems can be reduced to the following one. Let \mathfrak{T} be a set and let $(\Omega, \mathfrak{B}, P)$ be a complete probability space. Let χ defined on $\mathfrak{T} \times \Omega$ be a given \mathfrak{B} -measurable map into a linear metric space \mathfrak{S} and let $\Psi : \mathfrak{S} \rightarrow \mathbb{R}_+$ be a given *continuous sub-additive* functional.

Let $\Theta \subset \mathfrak{T}$ and suppose that $\forall \theta \in \Theta$ and $\forall z > 0$ one can find non-random $U(\theta, z) > 0$ and $c > 0$ such that

$$P \{ [\Psi(\chi_\theta) - U(\theta, z)] > 0 \} \leq ce^{-z}. \quad (1.1)$$

Assuming additionally that $\lambda U(\cdot, z) \geq U(\cdot, \lambda z)$ for any $z > 0$, $\lambda \geq 1$ we also have for any $q \geq 1$

$$E \left\{ [\Psi(\chi_\theta) - U(\theta, z)]_+^q \right\} \leq c\Gamma(q+1)[U(\theta, 1)]^q e^{-z}, \quad \forall z \geq 1, \quad (1.2)$$

where Γ is gamma-function and $[a]_+$ is the positive part of a .

The problem which we address now consists in a finding of $\mathbf{U}(\theta, z)$ and $\mathbf{U}_q(\theta, z)$ satisfying

$$\mathbb{P} \left\{ \sup_{\theta \in \Theta} [\Psi(\chi_\theta) - \mathbf{U}(\theta, z)] > 0 \right\} \leq \mathbf{c} e^{-z}; \quad \forall z \geq 1 \quad (1.3)$$

$$\mathbb{E} \left\{ \sup_{\theta \in \Theta} [\Psi(\chi_\theta) - \mathbf{U}_q(\theta, z)]_+ \right\}^q \leq \mathbf{c}_q \left[\inf_{\theta \in \Theta} \mathbf{U}_q(\theta, 1) \right]^q e^{-z}, \quad \forall z \geq 1, \quad (1.4)$$

where \mathbf{c} and \mathbf{c}_q are constants. If (1.3) and (1.4) hold we will say that $\mathbf{U}(\cdot, \cdot)$ and $\mathbf{U}_q(\cdot, \cdot)$ are *upper functions* for the collection of random variables $\{\Psi(\chi_\theta), \theta \in \Theta\}$.

The main questions which we would like to answer are the following.

- Do $\mathbf{U}(\cdot, \cdot)$ and $\mathbf{U}_q(\cdot, \cdot)$ coincide with $U(\cdot, \cdot)$ up to numerical constants or there is a "price to pay" for passing from pointwise results (1.1)–(1.2) to uniform ones given in (1.3)–(1.4)?
- Do $\mathbf{U}(\cdot, \cdot)$ and $\mathbf{U}_q(\cdot, \cdot)$ coincide up to numerical constants? In other words should one to pay the same price for the probability and moment's bounds?

We will show that a payment exists and in general $\mathbf{U}_q(\cdot, \cdot) \gg \mathbf{U}(\cdot, \cdot) \gg U(\cdot, \cdot)$. Thus, we will seek $\mathbf{U}(\cdot, \cdot)$ and $\mathbf{U}_q(\cdot, \cdot)$ satisfying (1.3) and (1.4) and "minimally" separated away from $U(\cdot, \cdot)$. We will realize this program under the following condition.

Assumption 1. 1. *There exist $A : \mathfrak{T} \rightarrow \mathbb{R}_+$, $B : \mathfrak{T} \rightarrow \mathbb{R}_+$ and $\mathbf{c} > 0$ such that $\forall z > 0$*

$$\mathbb{P} \{ \Psi(\chi_\theta) \geq z \} \leq \mathbf{c} \exp \left\{ -\frac{z^2}{A^2(\theta) + B(\theta)z} \right\}, \quad \forall \theta \in \Theta. \quad (1.5)$$

2. *There exist $\mathbf{a} : \mathfrak{T} \times \mathfrak{T} \rightarrow \mathbb{R}_+$ and $\mathbf{b} : \mathfrak{T} \times \mathfrak{T} \rightarrow \mathbb{R}_+$ such that $\forall z > 0$*

$$\mathbb{P} \{ \Psi(\chi_{\theta_1} - \chi_{\theta_2}) \geq z \} \leq \mathbf{c} \exp \left\{ -\frac{z^2}{\mathbf{a}^2(\theta_1, \theta_2) + \mathbf{b}(\theta_1, \theta_2)z} \right\}, \quad \forall \theta_1, \theta_2 \in \Theta. \quad (1.6)$$

Remark 1. *If Assumption 1 (1) holds on \mathfrak{T} (not only on Θ), \mathfrak{T} is linear space and if, additionally, the map χ_t is linear on \mathfrak{T} , then the Assumption 1 (2) is automatically fulfilled since one can take $\mathbf{a}(\mathbf{t}_1, \mathbf{t}_2) = A(\mathbf{t}_1 - \mathbf{t}_2)$ and $\mathbf{b}(\mathbf{t}_1, \mathbf{t}_2) = B(\mathbf{t}_1 - \mathbf{t}_2)$, $\mathbf{t}_1, \mathbf{t}_2 \in \mathfrak{T}$.*

Remark 2. *We can easily deduce from (1.5) that for any $\theta \in \Theta$*

$$\mathbb{P} \{ \Psi(\chi_\theta) \geq A(\theta)\sqrt{z} + B(\theta)z \} \leq \mathbf{c} \exp \{ -z \}, \quad \forall z \geq 0; \quad (1.7)$$

$$\mathbb{E} \left\{ \Psi(\chi_\theta) - [A(\theta)\sqrt{z} + B(\theta)z] \right\}_+^q \leq \mathbf{c} \Gamma(q+1) [A(\theta) + B(\theta)]^q \exp \{ -z \}, \quad \forall z \geq 1. \quad (1.8)$$

Therefore, (1.1)–(1.2) hold with $U(\theta, z) = A(\theta)\sqrt{z} + B(\theta)z$.

Assumption 1 is not new. In particular, it can be found in slightly different form in van der Vaart and Wellner (1996), Talagrand (2005), where this assumption is used for deriving the bound for $\mathbb{E} [\sup_{\theta \in \Theta} \Psi(\chi_\theta)]$. The usual technique is based on the *chaining argument* available in view of (1.6). It is worth mentioning that uniform probability and moment bounds for $[\sup_{\theta \in \Theta} \Psi(\chi_\theta)]$ in the case where χ_θ is empirical or gaussian process are a subject of vast literature, see, e.g., Alexander (1984), Talagrand (1994), Lifshits (1995), van der Vaart and Wellner (1996), van de Geer (2000), Massart (2000), Bousquet (2002), Giné and Koltchinskii (2006) among many others. Such bounds play an important role in establishing the laws of iterative logarithm and central limit theorems [see, e.g., Alexander (1984) and Giné and Zinn (1984)].

However much less attention was paid to finding of upper functions. The majority of the papers, where such problems are considered, contains asymptotical results, see, i.e. Kalinauskaitė (1966), Qualls and Watanabe (1972), Bobkov (1988), Shiryaev et al. (2002) and references therein. We would like especially mention the paper Egishyants and Ostrovskii (1996), where upper function satisfying the inequalities similar to (1.3), was obtained for the modulus of continuity of random fields satisfying the Cramer condition.

The researches carried out in the present paper complete the investigations done in Goldenshluger and Lepski (2011), where the upper functions as well as inequalities (1.3)–(1.4) were obtained under following condition: χ_t is linear and there are $A : \mathfrak{T} \rightarrow \mathbb{R}_+$, $B : \mathfrak{T} \rightarrow \mathbb{R}_+$, $V : \mathfrak{T} \rightarrow \mathbb{R}_+$ such that

$$\mathbb{P} \{ \Psi(\chi_t) - V(t) \geq z \} \leq g \left(\frac{z^2}{A^2(t) + B(t)z} \right), \quad \forall t \in \mathfrak{T},$$

where $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a strictly decreasing to zero function. We note that if $g(x) = e^{-x}$ and $V \equiv 0$ this assumption coincides with (1.5) and, since χ_t is linear (1.6) is automatically fulfilled, see Remark 1. In Goldenshluger and Lepski (2011) under additional assumption imposed on A, B, V and $\Theta \subset \mathfrak{T}$ the upper functions for the collection $\{\Psi(\chi_\theta), \theta \in \Theta\}$ were found. As it was shown that they coincide with the function V up to universal constants! The imposed assumptions do not admit the case $V \equiv 0$ that, as it was said above, leads to some "price to pay" for passing from pointwise results (1.1)–(1.2) to uniform ones given in (1.3)–(1.4).

To derive upper functions satisfying (1.3)–(1.4) we complete Assumption 1 by the following conditions.

Assumption 2. $\chi_\bullet : \mathfrak{T} \rightarrow \mathfrak{S}$ is continuous P-a.s.

Mappings a and b are semi-metrics on \mathfrak{T} and Θ is totally bounded with respect to $a \vee b$.

$$\overline{A}_\Theta := \sup_{\theta \in \Theta} A(\theta) < \infty, \quad \overline{B}_\Theta := \sup_{\theta \in \Theta} B(\theta) < \infty.$$

Denote by \mathbb{S} the following set of real functions:

$$\mathbb{S} = \left\{ s : \mathbb{R} \rightarrow \mathbb{R}_+ \setminus \{0\} : \sum_{k=0}^{\infty} s(2^{k/2}) \leq 1 \right\}.$$

For any $\tilde{\Theta} \subseteq \Theta$ and any semi-metric d on \mathfrak{T} let $\mathfrak{E}_{\tilde{\Theta}, d}(\delta)$, $\delta > 0$, denote the entropy of $\tilde{\Theta}$ measured in d . For any $x > 0$, $\tilde{\Theta} \subseteq \Theta$ and $s \in \mathbb{S}$ define the quantities

$$e_s^{(a)}(x, \tilde{\Theta}) = \sup_{\delta > 0} \delta^{-2} \mathfrak{E}_{\tilde{\Theta}, a} \left(x(48\delta)^{-1} s(\delta) \right), \quad e_s^{(b)}(x, \tilde{\Theta}) = \sup_{\delta > 0} \delta^{-1} \mathfrak{E}_{\tilde{\Theta}, b} \left(x(48\delta)^{-1} s(\delta) \right). \quad (1.9)$$

Assumption 3. There exist $s_1, s_2 \in \mathbb{S}$ such that $\forall x > 0$

$$e_{s_1}^{(a)}(x, \Theta) < \infty, \quad e_{s_2}^{(b)}(x, \Theta) < \infty.$$

What is this paper about? In the next section we construct upper functions for $\{\Psi(\chi_\theta), \theta \in \Theta\}$ and prove for them the inequalities (1.3)–(1.4) under Assumptions 1–3. We show that they are completely determined by the functions A and B and by the entropies of their level sets measured in semi-metrics a and b . We will see that obtained upper functions do not coincide with $U(\theta, z) = A(\theta)\sqrt{z} + B(\theta)z$, see Remark 2, and provide with explicit expression for the "price to be paid for uniformity". In particular, if $\underline{A} := \inf_{\theta \in \Theta} A(\theta) > 0$ and $\underline{B} := \inf_{\theta \in \Theta} B(\theta) > 0$, we prove that this "price" can be expressed as a given function of $A(\theta)/\underline{A}$ and $B(\theta)/\underline{B}$.

In this context it is interesting to compare our results with the usual probability bounds for $\sup_{\theta \in \Theta} \Psi(\chi_\theta)$ above $E\{\sup_{\theta \in \Theta} \Psi(\chi_\theta)\}$ obtained from Talagrand's or the Borel-Sudakov-Tsirelson inequality (when available), combined with uniform bounds for $E\{\sup_{\theta \in \Theta} \Psi(\chi_\theta)\}$ proved in Talagrand (2005), Theorem 1.2.7, under condition close to our Assumption 1. Following this strategy we will come to upper functions which are constant in θ . The main question is then what can one gain using the technique developed in the paper with respect to the aforementioned approach? We do not think that the answer can be done under "abstract considerations", i.e. under Assumptions 1-3 since it would require to prove that the found "price to be paid for uniformity" is minimal. However, in concrete examples it seems to be possible, in particular for some problems studied in mathematical statistics. Let us mention some of them. First, we recall that upper functions are used in all known constructions of adaptive procedures. Next, the use of upper functions being constant in θ will lead to adaptive estimators which are not optimal (remind, that the adaptive estimation theory is equipped with very developed criterion of optimality). Contrary to this, in all known to the author examples the use of upper functions found in Propositions 2 and 3 allows to construct optimal adaptive procedures (for more details see discussion after Proposition 2). Also, we would like to emphasize that upper functions being constant in θ and the probability bounds related to them are similar to the construction and the results described in Proposition 1, which is, in its turn, the initial step for our considerations. This step as well as the Talagrand's bounds are obtained from chaining argument under, for instance, Assumption 1. In some sense one of our goals is to show that the use of concentration inequalities (which cannot be guaranteed by a condition similar to Assumption 1) in the construction of upper functions is not necessary. In particular, in Section 3.1 we derive an upper function for the \mathbb{L}_p -norms of Wiener integrals and deduce the corresponding probability bounds directly from Proposition 3 without passing to the concentration inequalities.

As it was mentioned above upper functions for random objects appear in various areas of mathematical statistics. To apply them in the construction of statistical procedures they have to be computed explicitly. In particular the study of the adaptive estimation in the density model requires to find upper functions for the empirical processes of different kind. For the majority of existed problems Assumption 1 follows from the Bernstein's inequality. However the application of Propositions 2-3 requires to compute the functions \mathcal{E} or $\widehat{\mathcal{E}}$ (involved in the description of upper functions) and there is no a general recipe how to do it. One of our main objectives is to provide with rather general assumptions under which the latter quantities can be computed explicitly. In particular, we provide with the condition (main assumption in Part II of the paper) and to the best of our knowledge the assumptions of such kind have not been appeared in the existing literature. Under this assumption the upper functions are found for the variety of particular problems.

Organization of the paper. The paper is divided into 3 parts which are supposed to be published separately.

Part I. In Section 2 of we construct upper functions for $\{\Psi(\chi_\theta), \theta \in \Theta\}$ and prove for them the inequalities (1.3)–(1.4) under Assumptions 1–3. In fact we present two different constructions which will be referred to upper functions of the first and second type (Propositions 2 and 3). We also derive some consequences related to the upper functions for modulus of continuity of random real-valued mappings (Propositions 4 and 5). In Section 3 we apply Propositions 3 and 4 to gaussian random functions. In Section 3.1 we derive upper functions for \mathbb{L}_p -norm of some Wiener integrals (Theorem 1) and in Section 3.2 we study the local modulus of continuity of gaussian functions defined on a metric space satisfying doubling condition (Theorem 2). Proofs are given in Sections 4–5.

Part II and Part III are devoted to the detailed consideration of generalized empirical processes. We provide with rather general assumption under which the upper functions admit the explicit expression. We also establish non-asymptotical versions of the law of iterated logarithm and the law of logarithm. Then we apply the developed technique to empirical processes possessing some special structure.

2. General setting

Denote by $\mathcal{S}_{a,b}$ the subset of $\mathbb{S} \times \mathbb{S}$ for which Assumption 3 holds and let A, B, a and b be any mappings for which Assumption 1 is fulfilled.

For any $\vec{s} = (s_1, s_2) \in \mathcal{S}_{a,b}$, any $\varkappa = (\varkappa_1, \varkappa_2)$, $\varkappa_1 > 0, \varkappa_2 > 0$, and any $\tilde{\Theta} \subseteq \Theta$ put

$$e_{\vec{s}}(\varkappa, \tilde{\Theta}) = e_{s_1}^{(a)}(\varkappa_1, \tilde{\Theta}) + e_{s_2}^{(b)}(\varkappa_2, \tilde{\Theta}). \quad (2.1)$$

2.1. Inequalities for the suprema

Put for any $\tilde{\Theta} \subseteq \Theta$, any $\varepsilon > 0$ and any $y \geq 0$

$$U_{\vec{s}}^{(\varepsilon)}(y, \varkappa, \tilde{\Theta}) = \varkappa_1 \sqrt{2[1 + \varepsilon^{-1}]^2 e_{\vec{s}}(\varkappa, \tilde{\Theta}) + y} + \varkappa_2 (2[1 + \varepsilon^{-1}]^2 e_{\vec{s}}(\varkappa, \tilde{\Theta}) + y).$$

Proposition 1. *Let Assumptions 1-3 hold and let $\tilde{\Theta} \subseteq \Theta$ be fixed. Then for any $\tilde{\varkappa}$ satisfying $\tilde{\varkappa}_1 \geq \sup_{\theta \in \tilde{\Theta}} A(\theta)$ and $\tilde{\varkappa}_2 \geq \sup_{\theta \in \tilde{\Theta}} B(\theta)$, any $\vec{s} \in \mathcal{S}_{a,b}$, $\varepsilon \in (0, \sqrt{2} - 1]$ and $y \geq 1$,*

$$\mathbb{P} \left\{ \sup_{\theta \in \tilde{\Theta}} \Psi(\chi_\theta) \geq U_{\vec{s}}^{(\varepsilon)}(y, \tilde{\varkappa}, \tilde{\Theta}) \right\} \leq 2c \exp \left\{ -y/(1 + \varepsilon)^2 \right\}.$$

Moreover, for any $q \geq 1$

$$\mathbb{E} \left\{ \sup_{\theta \in \tilde{\Theta}} \Psi(\chi_\theta) - U_{\vec{s}}^{(\varepsilon)}(y, \tilde{\varkappa}, \tilde{\Theta}) \right\}_+^q \leq 2c\Gamma(q+1) \left[(1 + \varepsilon)^2 y^{-1} U_{\vec{s}}^{(\varepsilon)}(y, \tilde{\varkappa}, \tilde{\Theta}) \right]^q \exp \left\{ -y/(1 + \varepsilon)^2 \right\}.$$

We remark that $\sup_{\theta \in \tilde{\Theta}} \Psi(\chi_\theta)$ is \mathfrak{B} -measurable for any $\tilde{\Theta} \subseteq \Theta$ since Ψ is continuous, the mapping $\theta \mapsto \chi_\theta$ is continuous P-a.s., Θ is a totally bounded set and considered probability space is complete (see, e.g. Lemma 1 below).

Discussion We will see that the Proposition 1 is crucial technical tool for deriving upper functions. It contains the main ingredient of our future construction the quantity $e_{\vec{s}}$. The important issue in this context is the choice of $\vec{s} \in \mathcal{S}_{a,b}$. For many particular problems it is sufficient to choose $\vec{s} = (s^*, s^*)$, where

$$s^*(x) = (6/\pi^2)(1 + [\ln x]^2)^{-1}, \quad x \geq 0. \quad (2.2)$$

This choice is explained by two simple reasons: its explicit description allowing to compute the quantity $e_{\vec{s}}$ in particular problem and the logarithmical decay of this function when $x \rightarrow \infty$. In view of the latter remark we can consider the set Θ those entropy obeys the restriction which is closer to the minimal one (c.f. Sudakov lower bound for gaussian random functions Lifshits (1995)). We note, however, that there exist examples where \vec{s} has to be chosen on a more special way (see Theorem 1).

Let us now discuss the role of the parameter ε . In most particular problems considered in the paper we will not be interested in optimization of the numerical constants involved in the description of upper functions. If so, the choice of this parameter can be done in arbitrary way and we will put $\varepsilon = \sqrt{2} - 1$ to simplify the notations and computations. Note, however, that there are some problems (see, for instance Section 3.2), where ε must be chosen carefully. The typical requirements to this choice is $\varepsilon = \varepsilon(y)$ and

$$\varepsilon(y) \rightarrow 0, \quad y\varepsilon^2(y) \rightarrow 0, \quad y \rightarrow \infty.$$

The bounds similar to those presented in Proposition 1 are the subject of vast literature see, for instance, the books Lifshits (1995), van der Vaart and Wellner (1996) or van de Geer (2000). Note, however, that the results presented in the proposition may have an independent interest, at least, for the problems where the quantity $e_{\tilde{s}}$ can be expressed explicitly. In this case under rather general conditions it is possible, putting $\tilde{\Theta} = \Theta$ and $\tilde{\varepsilon} = (\bar{A}_{\Theta}, \bar{B}_{\Theta})$, to compute the tail probability as well as the expected value of the suprema of random mappings. Note also that Assumptions 1-3 guarantee that $E \{ \sup_{\theta \in \Theta} \Psi(\chi_{\theta}) \}^q$ is finite for any $q \geq 1$.

2.2. Upper functions of the first and second type

We will now use Proposition 1 in order to derive the upper functions for $\Psi(\chi_{\theta})$ on Θ . Denote $\underline{A} = \inf_{\theta \in \Theta} A(\theta)$ and $\underline{B} = \inf_{\theta \in \Theta} B(\theta)$.

We present two kinds of upper functions for $\Psi(\chi_{\theta})$ on Θ which we will refer to upper functions of the first and second type. The first construction is completely determined by the functions A , B and by the semi-metrics a and b . It requires however the additional condition $\underline{A} > 0, \underline{B} > 0$. We will use corresponding results for the particular problems in Parts II and III.

The second construction is related to some special structure imposed on the set Θ . Namely we will suppose that $\Theta = \cup_{\alpha} \Theta_{\alpha} \in \mathfrak{A}$, where $\{\Theta_{\alpha}, \alpha \in \mathfrak{A}\}$ is a given collection of sets. Here we will be interested in a finding of upper function for $\sup_{\theta \in \Theta_{\alpha}} \Psi(\chi_{\theta})$ on \mathfrak{A} , which can be also viewed as an upper function for $\Psi(\chi_{\theta})$ on Θ . The corresponding results are used in order to obtain rather precise inequalities for the modulus of continuity of random functions, Section 3.2. Moreover we apply this bound for deriving of an upper function for the \mathbb{L}_p -norms of Wiener integrals, Section 3.1. We deduce the corresponding inequality directly from Proposition 3 below without passing to the concentration inequalities.

We finish this short introduction with the following remark. In order to establish the inequalities (1.3)–(1.4) for the upper functions presented below we will need to prove that corresponding supremum is a random variable. The result below is sufficient for all problems considered in the paper and before we start the proofs we will not discuss the measurability issue.

Lemma 1. *Let \mathfrak{T} be the set equipped with the metric \mathfrak{d} , $(\Omega, \mathfrak{B}, \mathbf{P})$ be a complete probability space and $\zeta : \Omega \times \mathfrak{T} \rightarrow \mathbb{R}$ be \mathbf{P} -a.s. continuous. Let \mathfrak{Z} be a set, $g : \mathfrak{Z} \rightarrow \mathbb{R}$ be a given function and $\{\mathfrak{T}_{\mathfrak{z}} \subseteq \mathfrak{T}, \mathfrak{z} \in \mathfrak{Z}\}$ be an arbitrary sequence of sets. If \mathfrak{T} is totally bounded then $\sup_{\mathfrak{z} \in \mathfrak{Z}} \left[\sup_{t \in \mathfrak{T}_{\mathfrak{z}}} \zeta(t, \cdot) - g(\mathfrak{z}) \right]$ is \mathfrak{B} -measurable.*

The proof of the lemma is given in Appendix. We would like to emphasize that there is no any assumption imposed on the function g , index set \mathfrak{Z} and on the collection $\{\mathfrak{T}_{\mathfrak{z}} \subseteq \mathfrak{T}, \mathfrak{z} \in \mathfrak{Z}\}$.

Putting $\mathfrak{Z} = \mathfrak{T}$ and $\mathfrak{T}_{\mathfrak{t}} = \{\mathfrak{t}\}$ we come to the following consequence of Lemma 1.

Corollary 1. *Under assumptions of Lemma 1 $\sup_{t \in \mathfrak{T}} [\zeta(t, \cdot) - g(t)]$ is \mathfrak{B} -measurable.*

Upper functions of the first type As it was said above throughout this section we will suppose that $\underline{A} > 0, \underline{B} > 0$. Put for any $t > 0$

$$\Theta_A(t) = \{\theta \in \Theta : A(\theta) \leq t\}, \quad \Theta_B(t) = \{\theta \in \Theta : B(\theta) \leq t\}.$$

For any $\vec{s} \in \mathcal{S}_{a,b}$ introduce the function

$$\mathcal{E}_{\vec{s}}(u, v) = e_{s_1}^{(a)}(\underline{A}u, \Theta_A(\underline{A}u)) + e_{s_2}^{(b)}(\underline{B}v, \Theta_B(\underline{B}v)), \quad u, v \geq 1. \quad (2.3)$$

Denote also $\ell(u) = \ln\{1 + \ln(u)\} + 2 \ln\{1 + \ln\{1 + \ln(u)\}\}$ and set for any θ and $\varepsilon > 0, r \geq 0$

$$P_\varepsilon(\theta) = 2[1 + \varepsilon^{-1}]^2 \mathcal{E}_{\vec{s}}(\mathcal{A}_\varepsilon(\theta), \mathcal{B}_\varepsilon(\theta)) + (1 + \varepsilon)^2 [\ell(\mathcal{A}_\varepsilon(\theta)) + \ell(\mathcal{B}_\varepsilon(\theta))]; \quad (2.4)$$

$$M_{\varepsilon,r}(\theta) = (1 + \varepsilon)^2 \left\{ 2[1 + \varepsilon^{-1}]^2 \mathcal{E}_{\vec{s}}(\mathcal{A}_\varepsilon(\theta), \mathcal{B}_\varepsilon(\theta)) + (\varepsilon + r) \ln [\mathcal{A}_\varepsilon(\theta) \mathcal{B}_\varepsilon(\theta)] \right\}, \quad (2.5)$$

where $\mathcal{A}_\varepsilon(\theta) = (1 + \varepsilon)[A(\theta)/\underline{A}]$ and $\mathcal{B}_\varepsilon(\theta) = (1 + \varepsilon)[B(\theta)/\underline{B}]$. Define for any $z \geq 0$

$$V^{(z,\varepsilon)}(\theta) = (1 + \varepsilon)^2 \left(A(\theta) \sqrt{P_\varepsilon(\theta) + (1 + \varepsilon)^2 z} + B(\theta) [P_\varepsilon(\theta) + (1 + \varepsilon)^2 z] \right); \quad (2.6)$$

$$U^{(z,\varepsilon,r)}(\theta) = (1 + \varepsilon)^2 \left(A(\theta) \sqrt{M_{\varepsilon,r}(\theta) + (1 + \varepsilon)^2 z} + B(\theta) [M_{\varepsilon,r}(\theta) + (1 + \varepsilon)^2 z] \right). \quad (2.7)$$

In the proposition below we prove that the functions defined in (2.6) and (2.7) are upper functions for $\Psi(\chi_\theta)$ on Θ . We remark that they are completely determined by the functions A and B and by the entropies of their level sets measured in semi-metrics a and b . The number ε and the couple of functions \vec{s} can be viewed as tuning parameters allowing either to weaken assumptions or to obtain sharper bounds but they are not related to the random functional $\Psi(\chi_\theta)$ itself.

Proposition 2. *Let Assumptions 1-3 be fulfilled. Then $\forall \vec{s} \in \mathcal{S}_{a,b}, \forall \varepsilon \in (0, \sqrt{2} - 1]$ and $\forall z \geq 1$*

$$\begin{aligned} \mathbb{P} \left\{ \sup_{\theta \in \Theta} [\Psi(\chi_\theta) - V^{(z,\varepsilon)}(\theta)] \geq 0 \right\} &\leq 2c \left[1 + \left[\ln\{1 + \ln(1 + \varepsilon)\} \right]^{-2} \right]^2 \exp\{-z\}; \\ \mathbb{E} \left\{ \sup_{\theta \in \Theta} [\Psi(\chi_\theta) - U^{(z,\varepsilon,r)}(\theta)] \right\}_+^q &\leq c 2^{(5q/2)+2} \Gamma(q+1) \varepsilon^{-q-4} [\underline{A} \vee \underline{B}]^q \exp\{-z\}. \end{aligned}$$

It is obvious that the assertions of the proposition remain valid if one replaces the function $\mathcal{E}_{\vec{s}}$ by any its upper bound. It is important since the exact computation of this function is too complicated in general. We note that the role of the latter function in our construction is similar to those which Dudley integral plays in the computations of the expectation of the suprema of gaussian or sub-gaussian processes Lifshits (1995), Talagrand (2005).

Price to pay for "uniformity" We remark that in view of (1.7) and (1.8), the function $U^{(z)}(\theta) := A(\theta)\sqrt{z} + B(\theta)z$ can be viewed as "pointwise upper function" for $\Psi(\chi_\theta)$, i.e. for fixed θ . Comparing the inequalities (1.7) and (1.8) with those given in Proposition 2 we conclude that they differ from each other by numerical constants only. In this context, the functions $P_\varepsilon(\cdot)$ and $M_{\varepsilon,r}(\cdot)$ given by (2.4) and (2.5) can be viewed as price to pay for "uniformity". That means that in order to pass from "pointwise" result to the "uniform" one we need, roughly speaking, to multiply $A(\cdot)$

by $\sqrt{P_\varepsilon(\cdot)}$ or $\sqrt{M_{\varepsilon,r}(\cdot)}$ and $B(\cdot)$ by $P_\varepsilon(\cdot)$ or $M_{\varepsilon,r}(\cdot)$. The question, arising naturally: is such payment necessary or minimal? In this context it is worth mentioning the relation between well-known phenomenon in adaptive estimation, called *price to pay for adaptation* Lepski (1991), Lepski and Spokoiny (1997) and Spokoiny (1996), and what we call here *price to pay for uniformity*. We have no place here to describe this relation in detail and mention only several facts.

First let us remark almost all constructions of adaptive estimators (model selection Barron et al. (1999), risk hull minimization Cavalier and Golubev (2006), Lepski method Lepski (1991), or recently developed universal estimation routine Goldenshluger and Lepski (2008, 2009)) involve the upper functions for stochastic objects of different kinds. Next, it is known that there are two types of *price to pay for adaptation*: (ln)-price, Lepski (1991) and (ln ln)-price, Spokoiny (1996). The (ln)-price appears in the problems where the risk of estimation procedures is described by a power loss-functions and it corresponds to the function $M_{\varepsilon,r}(\cdot)$, where the parameter r is a power. The (ln ln)-price appears in the case of bounded losses that corresponds to the function $P_\varepsilon(\cdot)$. Since the theory of adaptive estimation is equipped with very developed criteria of optimality, Lepski (1991), Tsybakov (1998), Kluchnikoff (2005), we might assert that the payment for uniformity is optimal if the use of corresponding upper function leads to optimal adaptive estimators.

We finish the discussion concerning the statements of Proposition 2 with the following remark. Comparing the result given in (1.8) with the second assertion of Proposition 2 we can state that the inequality obtained there is very precise since, remind, $\underline{A} = \inf_{\theta \in \Theta} A(\theta)$ and $\underline{B} = \inf_{\theta \in \Theta} B(\theta)$.

Upper functions of the second type Suppose that we are given by the collection $\{\Theta_\alpha, \alpha \in \mathfrak{A}\}$, satisfying $\Theta = \cup_{\alpha \in \mathfrak{A}} \Theta_\alpha$, and by two mappings $\tau_1 : \mathfrak{A} \rightarrow (0, \bar{\tau}_1]$, $\tau_2 : \mathfrak{A} \rightarrow (0, \bar{\tau}_2]$, where $\bar{\tau}_1, \bar{\tau}_2 < \infty$. For any $u > 0$ put

$$\begin{aligned} \Theta'_1(u) &= \bigcup_{\alpha: \tau_1(\alpha) \leq u} \Theta_\alpha, & g_A^*(u) &= \sup_{\theta \in \Theta'_1(u)} A(\theta); \\ \Theta'_2(u) &= \bigcup_{\alpha: \tau_2(\alpha) \leq u} \Theta_\alpha, & g_B^*(u) &= \sup_{\theta \in \Theta'_2(u)} B(\theta), \end{aligned}$$

and let g_A and g_B be arbitrary chosen increasing functions, satisfying $g_A \geq g_A^*$ and $g_B \geq g_B^*$ (we note that obviously g_A^* and g_B^* are increasing).

Since $\Theta'_1(\cdot), \Theta'_2(\cdot) \subseteq \Theta$, in view of Assumption 3 for any $u, v > 0$ one can find the functions $s_1(u, \cdot)$ and $s_2(v, \cdot)$ for which the latter assumption is fulfilled on $\Theta'_1(u)$ and $\Theta'_2(v)$ respectively. Let us suppose additionally that

$$\lambda_1 := \sup_{t \in [1, \sqrt{2}]} \sup_{x > \tau_1} \sup_{\delta > 0} \frac{s_1(xt, \delta)}{s_1(x, \delta)} < \infty, \quad \lambda_2 := \sup_{t \in [1, \sqrt{2}]} \sup_{x > \tau_2} \sup_{\delta > 0} \frac{s_2(xt, \delta)}{s_2(x, \delta)} < \infty, \quad (2.8)$$

where $\tau_1 = \inf_{\alpha} \tau_1(\alpha)$ and $\tau_2 = \inf_{\alpha} \tau_2(\alpha)$.

We remark that if the functions $s_1(u, \cdot)$ and $s_2(v, \cdot)$ are chosen independently of u, v then $\lambda_1 = \lambda_2 = 1$. It is also obvious that $\lambda_1, \lambda_2 \geq 1$.

The condition (2.8) allows us to define the function:

$$\mathcal{E}'(u, v) = e_{s_1(u, \cdot)}^{(a)} \left(\lambda_1^{-1} g_A(u), \Theta'_1(u) \right) + e_{s_2(v, \cdot)}^{(b)} \left(\lambda_2^{-1} g_B(v), \Theta'_2(v) \right), \quad u, v > 0. \quad (2.9)$$

We note that the function \mathcal{E}' is constructed similarly to the function \mathcal{E} used in the previous section, but now the functions s_1 and s_2 can be chosen in accordance with considered level sets.

At last, for any $\alpha \in \mathfrak{A}$ and any $\varepsilon > 0$ set

$$\widehat{\mathcal{E}}^{(\varepsilon)}(\alpha) = \mathcal{E}'\left((1 + \varepsilon)\tau_1(\alpha), (1 + \varepsilon)\tau_2(\alpha)\right).$$

Put $\delta_j = (1 + \varepsilon)^{-j}$, $j \geq 0$, and let $R_r : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$, $r \geq 0$, be an arbitrary family of *increasing* (or *decreasing*) in both arguments functions, satisfying for any $\varepsilon \in (0, \sqrt{2} - 1]$

$$\sum_{j=0}^J \sum_{k=0}^K [g_A(\bar{\tau}_1 \delta_j) \vee g_B(\bar{\tau}_2 \delta_k)]^r \exp \{-R_r(\bar{\tau}_1 \delta_j, \bar{\tau}_2 \delta_k)\} =: \mathcal{R}^{(\varepsilon, r)} < \infty. \quad (2.10)$$

Here integers J, K are defined as follows.

$$J = \lfloor \ln_{1+\varepsilon}(\bar{\tau}_1/\tau_1) \rfloor + 1, \quad K = \lfloor \ln_{1+\varepsilon}(\bar{\tau}_2/\tau_2) \rfloor + 1.$$

If $\tau_i = 0$, $i = 1, 2$, the corresponding quantity is put equal to infinity.

Set $\widehat{R}_r^{(\varepsilon)}(\alpha) = R_r(r_\varepsilon \tau_1(\alpha), r_\varepsilon \tau_2(\alpha))$, where $r_\varepsilon = (1 + \varepsilon)$ if R_r is increasing and $r_\varepsilon = (1 + \varepsilon)^{-1}$ if R_r is decreasing, and define

$$\begin{aligned} \widehat{U}^{(z, \varepsilon, r)}(\alpha) &= (1 + \varepsilon)g_A\left([1 + \varepsilon]^2 \tau_1(\alpha)\right) \sqrt{2[1 + \varepsilon^{-1}]^2 \widehat{\mathcal{E}}^{(\varepsilon)}(\alpha) + \widehat{R}_r^{(\varepsilon)}(\alpha) + z} \\ &+ (1 + \varepsilon)^2 g_B\left([1 + \varepsilon]^2 \tau_2(\alpha)\right) \left(2[1 + \varepsilon^{-1}]^2 \widehat{\mathcal{E}}^{(\varepsilon)}(\alpha) + \widehat{R}_r^{(\varepsilon)}(\alpha) + z\right). \end{aligned}$$

Below we assert that $\widehat{U}^{(z, \varepsilon, r)}$, $r = 0, r = q$, are upper functions for $\left[\sup_{\theta \in \Theta_\alpha} \Psi(\chi_\theta)\right]$ on \mathfrak{A} . However, before to present exact statements, let us briefly discuss some possible choices of the functions R_r . We would like to emphasize that the opportunity to select these functions allows to obtain quite different and precise results. First possible choice is given by

$$R_0(u, v) = \ell(\bar{\tau}_1 u^{-1}) + \ell(\bar{\tau}_2 v^{-1}), \quad R_r(u, v) = \varepsilon \left[\ln(\bar{\tau}_1 u^{-1}) + \ln(\bar{\tau}_2 v^{-1}) \right], \quad r > 0. \quad (2.11)$$

These functions are used in the problems in which $\widehat{\mathcal{E}}^{(\varepsilon)}(\cdot)$ is bounded by some absolute constant independent of all quantities involved in the description of the problem, assumptions etc.

This choice leads to the following values of the constants in (2.10):

$$\mathcal{R}^{(\varepsilon, 0)} \leq \left[2 + \left[\ln\{1 + \ln(1 + \varepsilon)\}\right]^{-2}\right]^2, \quad \mathcal{R}^{(\varepsilon, r)} < 4 \left[g_A(\bar{\tau}_1) \vee g_B(\bar{\tau}_2)\right]^r \varepsilon^{-4}. \quad (2.12)$$

Another important choice is given by $R_r = \mathcal{E}'$ independently of r , see, for instance, Theorem 1. In view of (2.10), this choice corresponds to the case when the function \mathcal{E}' increases to infinity.

Proposition 3. *Let Assumptions 1-3 be fulfilled. Then for any s_1, s_2 satisfying (2.8) and any R_r , $r \geq 0$, satisfying (2.10), for any $\varepsilon \in (0, \sqrt{2} - 1]$ and any $z \geq 1, q \geq 1$*

$$\begin{aligned} \mathbb{P} \left\{ \sup_{\alpha \in \mathfrak{A}} \left[\sup_{\theta \in \Theta_\alpha} \Psi(\chi_\theta) - \widehat{U}^{(z, \varepsilon, 0)}(\alpha) \right] \geq 0 \right\} &\leq 2c \mathcal{R}^{(\varepsilon, 0)} \exp\{-z\}; \\ \mathbb{E} \left\{ \sup_{\alpha \in \mathfrak{A}} \left[\sup_{\theta \in \Theta_\alpha} \Psi(\chi_\theta) - \widehat{U}^{(z, \varepsilon, q)}(\alpha) \right] \right\}_+^q &\leq c 2^{(5q/2)+1} \Gamma(q+1) \mathcal{R}^{(\varepsilon, q)} \varepsilon^{-q} \exp\{-z\}. \end{aligned}$$

Remark 3. *We note that the results of the proposition is very general. Indeed, there are no assumptions imposed on the collection Θ_α , $\alpha \in \mathfrak{A}$, and the functions τ_1, τ_2 can be chosen arbitrary. Moreover, the condition (2.10) is very mild, so the choice of functions R_r is quite flexible.*

2.3. Upper functions for the modulus of continuity of random mappings

In this section we apply Proposition 3 in order to derive upper functions for the local and global modulus of continuity of real-valued random mappings. It is worth mentioning that in this circle of problems the upper functions are actively exploited, see e.g. Egishyants and Ostrovskii (1996) and the references therein. We will suppose that Assumption 1 (2), Assumption 2 and Assumption 3 are verified, χ_t is real-valued random mapping defined on the metric space \mathfrak{T} , d is a semi-metric on \mathfrak{T} and $\Psi(\cdot) = |\cdot|$.

Upper function for local modulus of continuity Let θ_0 be a fixed element of Θ and set for any $\Delta \in (0, D_d(\Theta)]$, where $D_d(\Theta)$ is the diameter of Θ measured in the semi-metric d ,

$$m_\Delta(\theta_0) = \sup_{\theta \in \Theta_\Delta} |\chi_\theta - \chi_{\theta_0}|, \quad \Theta_\Delta = \{\theta \in \Theta : d(\theta, \theta_0) \leq \Delta\}.$$

Thus, $m_\Delta(\theta_0)$, $\Delta \in (0, D_d(\Theta)]$, is the local modulus of continuity of χ_θ in θ_0 measured in d .

If we put $\tilde{\chi}_\theta = \chi_\theta - \chi_{\theta_0}$, $\theta \in \Theta$, we assert first that Assumption 1 (2) can be viewed as Assumption 1 (1) for $\tilde{\chi}_\theta$ on Θ with $A(\cdot) = a(\cdot, \theta_0)$ and $B(\cdot) = b(\cdot, \theta_0)$. Next, noting that $\tilde{\chi}_{\theta_1} - \tilde{\chi}_{\theta_2} = \chi_{\theta_1} - \chi_{\theta_2}$ for any $\theta_1, \theta_2 \in \Theta$ we conclude that Assumption 1 (2) is verified for $\tilde{\chi}_\theta$ on Θ with a and b .

Thus, we can apply Proposition 3 with $\alpha = \Delta$, $\Theta_\alpha = \Theta_\Delta$, $\mathfrak{A} = (0, D_d(\Theta)]$ and we choose $\tau_1(\Delta) = \tau_2(\Delta) = \Delta$. This choice implies obviously for any $u \leq D_d(\Theta)$

$$\Theta'_1(u) = \Theta'_2(u) = \Theta_u, \quad g_A(u) = \sup_{\theta: d(\theta, \theta_0) \leq u} a(\theta, \theta_0), \quad g_B(u) = \sup_{\theta: d(\theta, \theta_0) \leq u} b(\theta, \theta_0).$$

Fix $\vec{s} \in \mathcal{S}_{a,b}$ and put for any $\Delta \in (0, D_d(\Theta)]$ and any $\varepsilon \in (0, \sqrt{2} - 1]$

$$\hat{\mathcal{E}}^{(\varepsilon)}(\Delta, \theta_0) = e_{s_1}^{(a)}(g_A([1 + \varepsilon]\Delta), \Theta_{[1 + \varepsilon]\Delta}) + e_{s_2}^{(b)}(g_B([1 + \varepsilon]\Delta), \Theta_{[1 + \varepsilon]\Delta}).$$

Here $e_{s_1}^{(a)}$ and $e_{s_2}^{(b)}$ are defined by (1.9). We also set $\lambda_1 = \lambda_2 = 1$ since the functions s_1, s_2 are chosen independently of the collection $\{\Theta_\Delta, \Delta \in (0, D_d(\Theta)]\}$.

Choose also $R_0(u, v) = \ell(\bar{\tau}_1 u^{-1}) + \ell(\bar{\tau}_2 v^{-1})$ and define

$$\begin{aligned} \hat{V}_{\vec{s}}^{(z, \varepsilon)}(\Delta, \theta_0) &= (1 + \varepsilon)g_A([1 + \varepsilon]^2 \Delta) \sqrt{2[1 + \varepsilon^{-1}]^2 \hat{\mathcal{E}}^{(\varepsilon)}(\Delta, \theta_0) + 2\ell((1 + \varepsilon)D_d(\Theta)/\Delta)} + z \\ &+ (1 + \varepsilon)^2 g_B([1 + \varepsilon]^2 \Delta) \left\{ 2[1 + \varepsilon^{-1}]^2 \hat{\mathcal{E}}^{(\varepsilon)}(\Delta, \theta_0) + 2\ell((1 + \varepsilon)D_d(\Theta)/\Delta) + z \right\}. \end{aligned}$$

Then, applying Proposition 3 and taking into account (2.12) we come to the following result.

Proposition 4. *Let Assumptions 1-3 be fulfilled. Then for any $\vec{s} \in \mathcal{S}_{a,b}$, $\varepsilon \in (0, \sqrt{2} - 1]$ and $z \geq 1$*

$$P \left\{ \sup_{\Delta \in (0, D_d(\Theta)]} [m_\Delta - \hat{V}_{\vec{s}}^{(z, \varepsilon)}(\Delta, \theta_0)] \geq 0 \right\} \leq 2c \left[2 + [\ln \{1 + \ln(1 + \varepsilon)\}]^{-2} \right]^2 \exp \{-z\}.$$

In Section 3 we apply Proposition 4 to gaussian random functions defined on a metric space satisfying so-called *doubling condition*.

Remark 4. If $b \equiv 0$, $d = a$ and $\sup_{\Delta \in (0, D_d(\Theta)]} \widehat{\mathcal{E}}^{(\varepsilon)}(\Delta, \theta_0) =: \widehat{\mathcal{E}}^{(\varepsilon)}(\theta_0) < \infty$, the upper function $\widehat{V}_s^{(z, \varepsilon)}$ has a very simple form

$$\widehat{V}_s^{(z, \varepsilon)}(\Delta, \theta_0) = (1 + \varepsilon)^3 \Delta \sqrt{2[1 + \varepsilon^{-1}]^2 \widehat{\mathcal{E}}^{(\varepsilon)}(\theta_0) + \ell((1 + \varepsilon)D_d(\Theta)/\Delta)} + z. \quad (2.13)$$

Hence, the result of Proposition 4 can be viewed as the non-asymptotical version of the law of iterated logarithm for sub-gaussian processes defined on some totally bounded subset of metric space. In this context it is worth mentioning the paper Egishyants and Ostrovskii (1996) where the upper functions for local and global modulus of continuity were found for the stochastic processes satisfying Cramer's condition.

Remark 5. We also note that we replaced in (2.13) the factor $2\ell(D_d(\Theta)/\Delta)$ appeared in the upper function used in Proposition 4 by $\ell(D_d(\Theta)/\Delta)$. It is explained by the fact that $\bar{\tau}_2 = 0$ in (2.11) since $B, b \equiv 0$. By the same reason, the probability bound in this case is given by $2c \left[2 + \left[\ln \{1 + \ln(1 + \varepsilon)\} \right]^{-2} \right] \exp \{-z\}$.

Upper function for global modulus of continuity Set $\Theta^{(2)} = \Theta \times \Theta$ and let for any $\vartheta = (\theta_1, \theta_2) \in \Theta^{(2)}$ and any $\Delta \in (0, D_d(\Theta)]$,

$$\zeta(\vartheta) = \chi_{\theta_1} - \chi_{\theta_2}, \quad \mathbf{m}_\Delta = \sup_{\vartheta \in \Theta^{(2)}} |\zeta_\vartheta|, \quad \Theta_\Delta^{(2)} = \{\vartheta \in \Theta^{(2)} : d(\theta_1, \theta_2) \leq \Delta\}.$$

Thus, \mathbf{m}_Δ , $\Delta \in (0, D_d(\Theta)]$, is the global modulus of continuity of χ_θ on Θ measured in d .

Put $\mathbf{A}(\vartheta) = a(\theta_1, \theta_2)$, $\mathbf{B}(\vartheta) = b(\theta_1, \theta_2)$, $\vartheta = (\theta_1, \theta_2) \in \Theta^{(2)}$, and equip $\Theta^{(2)}$ with the following semi-metrics: $\vartheta = (\theta_1, \theta_2)$, $\varsigma = (\varsigma_1, \varsigma_2) \in \Theta^{(2)}$

$$a^{(2)}(\vartheta, \varsigma) = 2[a(\theta_1, \varsigma_1) \vee a(\theta_2, \varsigma_2)], \quad b^{(2)}(\vartheta, \varsigma) = 2[b(\theta_1, \varsigma_1) \vee b(\theta_2, \varsigma_2)].$$

Some remarks are in order. We note first that Assumption 1 (2) can be viewed as Assumption 1 (1) for $\zeta(\vartheta)$ on $\Theta^{(2)}$ with $A = \mathbf{A}$ and $B = \mathbf{B}$.

Next we obtain in view of Assumption 1 (2) $\forall \vartheta, \varsigma \in \Theta^{(2)}$ and $\forall z > 0$

$$\begin{aligned} & \mathbb{P}\{|\zeta(\vartheta) - \zeta(\varsigma)| \geq z\} \leq \mathbb{P}\{|\chi_{\theta_1} - \chi_{\varsigma_1}| \geq z/2\} + \mathbb{P}\{|\chi_{\theta_2} - \chi_{\varsigma_2}| \geq z/2\} \\ & \leq c \exp \left\{ -\frac{z^2}{4[a(\theta_1, \varsigma_1)]^2 + 2b(\theta_1, \varsigma_1)z} \right\} + c \exp \left\{ -\frac{z^2}{4[a(\theta_2, \varsigma_2)]^2 + 2b(\theta_2, \varsigma_2)z} \right\} \\ & \leq c^{(2)} \exp \left\{ -\frac{z^2}{[a^{(2)}(\vartheta, \varsigma)]^2 + b^{(2)}(\vartheta, \varsigma)z} \right\} \end{aligned}$$

We conclude that Assumption 1 (2) holds for $\zeta(\vartheta)$ on $\Theta^{(2)}$ with $a = a^{(2)}$, $b = b^{(2)}$ and $c^{(2)} = 2c$.

Since obviously

$$\mathfrak{E}_{a^{(2)}, \Theta^{(2)}}(\varsigma) \leq 2\mathfrak{E}_{a, \Theta}(\varsigma/2), \quad \mathfrak{E}_{b^{(2)}, \Theta^{(2)}}(\varsigma) \leq 2\mathfrak{E}_{b, \Theta}(\varsigma/2), \quad \varsigma > 0, \quad (2.14)$$

we assert that Assumptions 2 and 3 are fulfilled on $\Theta^{(2)}$ with $a = a^{(2)}$ and $b = b^{(2)}$.

Put $\overline{\Theta}^{(2)} = \cup_{\Delta > 0} \Theta_{\Delta}^{(2)}$. Since $\overline{\Theta}^{(2)} \subset \Theta^{(2)}$ we can apply Proposition 3 with $\alpha = \Delta$, $\Theta_{\alpha} = \Theta_{\Delta}^{(2)}$, $\mathfrak{A} = (0, D_d(\Theta)]$ and we choose $\tau_1(\Delta) = \tau_2(\Delta) = \Delta$.

The latter choice implies obviously for any $u \leq D_d(\Theta)$

$$\Theta'_1(u) = \Theta'_2(u) = \Theta_u^{(2)}, \quad g_{\mathbf{A}}(u) = \sup_{\vartheta \in \Theta_u^{(2)}} \mathbf{A}(\vartheta), \quad g_{\mathbf{B}}(u) = \sup_{\vartheta \in \Theta_u^{(2)}} \mathbf{B}(\vartheta).$$

Fix $\vec{s} \in \mathcal{S}_{a,b}$ and put for any $\Delta \in (0, D_d(\Theta)]$ and any $\varepsilon \in (0, \sqrt{2} - 1]$

$$\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta) = e_{s_1}^{(a^{(2)})} \left(g_{\mathbf{A}}([1 + \varepsilon]\Delta), \Theta_{[1 + \varepsilon]\Delta}^{(2)} \right) + e_{s_2}^{(b^{(2)})} \left(g_{\mathbf{B}}([1 + \varepsilon]\Delta), \Theta_{[1 + \varepsilon]\Delta}^{(2)} \right).$$

Here $e_{s_1}^{(a^{(2)})}$ and $e_{s_2}^{(b^{(2)})}$ are defined by (1.9), where a, b are replaced by $a^{(2)}$ and $b^{(2)}$ respectively. We also set $\lambda_1 = \lambda_2 = 1$ since the functions s_1, s_2 are chosen independently of the collection $\{\Theta_{\Delta}^{(2)}, \Delta \in (0, D_d(\Theta)]\}$. Choose $R_0(u, v) = \ell(\bar{\tau}_1 u^{-1}) + \ell(\bar{\tau}_2 v^{-1})$ and define

$$\begin{aligned} \widehat{V}_{\vec{s}}^{(z, \varepsilon)}(\Delta) &= (1 + \varepsilon) g_{\mathbf{A}}([1 + \varepsilon]^2 \Delta) \sqrt{2[1 + \varepsilon^{-1}]^2 \widehat{\mathcal{E}}^{(\varepsilon)}(\Delta) + 2\ell((1 + \varepsilon)D_d(\Theta)/\Delta)} + z \\ &+ (1 + \varepsilon)^2 g_{\mathbf{B}}([1 + \varepsilon]^2 \Delta) \left\{ 2[1 + \varepsilon^{-1}]^2 \widehat{\mathcal{E}}^{(\varepsilon)}(\Delta) + 2\ell((1 + \varepsilon)D_d(\Theta)/\Delta) + z \right\}. \end{aligned}$$

Then, applying Proposition 3 and taking into account (2.12) we come to the following result.

Proposition 5. *Let Assumptions 1-3 be fulfilled. Then for any $\vec{s} \in \mathcal{S}_{a,b}$, $\varepsilon \in (0, \sqrt{2} - 1]$ and $z \geq 1$*

$$\mathbb{P} \left\{ \sup_{\Delta \in (0, D_d(\Theta)]} [\mathfrak{m}_{\Delta} - \widehat{V}_{\vec{s}}^{(z, \varepsilon)}(\Delta)] \geq 0 \right\} \leq 4c \left[2 + \left[\ln \{1 + \ln(1 + \varepsilon)\} \right]^{-2} \right]^2 \exp \{-z\}.$$

The obtained inequality allows, in particular, to prove that the families of probabilities measures generated by χ_{θ} is dense. This, in its turn, is crucial step in proving of the weak convergence of probabilities measures.

3. Gaussian random functions

In this section we apply Propositions 2-4 to the family of zero-mean gaussian random functions. Thus, let $\chi_{\theta}, \theta \in \Theta$, is a real valued continuous gaussian random function such that $E\chi_{\theta} = 0$, $\forall \theta \in \Theta$. We are interested first in finding an upper function for $|\chi_{\theta}|, \theta \in \Theta$. Let

$$V(\theta) = \sqrt{E|\chi_{\theta}|^2}, \quad \rho(\theta_1, \theta_2) = \sqrt{E|\chi_{\theta_1} - \chi_{\theta_2}|^2}$$

We remark that Assumption 1 holds with $c = 2$, $B \equiv 0$ and $b \equiv 0$ and $\forall A \geq \sqrt{2}V, \forall a \geq \sqrt{2}\rho$. Since $b \equiv 0$ Assumption 3 is reduced to

Assumption 3 [Gaussian case]. *There exist $s \in \mathbb{S}$ such that for any $x > 0$*

$$\sup_{\delta > 0} \delta^{-2} \mathfrak{E}_{\Theta, a}^{\sim} \left(x(48\delta)^{-1} s(\delta) \right) < \infty.$$

Thus, if the latter assumption holds, Propositions 2-4 can be applied.

The aim of this section is to find uppers functions for quite different functionals of various gaussian processes. We would like to emphasize that the original problem is not always related to the consideration of $|\chi_{\theta}|, \theta \in \Theta$, although such problems are also studied. The idea is to reduce it (if necessary) to those for which one of Propositions 2-4 can be used. Without special mentioning we will always consider a separable modification of $\chi_{\theta}, \theta \in \Theta$.

3.1. Upper functions for \mathbb{L}_p -norms of Wiener integrals

Let $K : \mathbb{R}^d \rightarrow \mathbb{R}$ be a continuous compactly supported function such that $\|K\|_\infty < \infty$. Without loss of generality we will assume that the support of K is $[-1/2, 1/2]^d$. Let $0 < h^{(\min)} \leq h^{(\max)} \leq 1$ be given numbers. Put $\mathcal{H} = [h^{(\min)}, h^{(\max)}]$ and let $K_h(u) = h^{-d}K(u_1/h, \dots, u_d/h)$, $h \in \mathcal{H}$, $u \in \mathbb{R}^d$. Let $b(dt)$ is white noise on \mathbb{R}^d and consider the family of gaussian random fields

$$\xi_h(t) = \int_{\mathbb{R}^d} K_h(t - u) b(du), \quad h \in \mathcal{H}.$$

Let $\mathbb{K}_\mu = [-\mu/2, \mu/2]^d$, $\mu \geq 1$, be a given cube and let for any $1 \leq p < \infty$

$$\|\xi_h\|_p = \left(\int_{\mathbb{K}_\mu} |\xi_h(t)|^p dt \right)^{\frac{1}{p}}.$$

The objective is to find an upper function for $\|\xi_h\|_p$ on \mathcal{H} and later on C_1, C_2, \dots , denote the constants completely determined by d, p, μ, γ and K . It is worth mentioning that the explicit values of these constants can be found and some of them are given in the proof of the theorem.

We will be interested only the case $2 \leq p < \infty$, since for any $p \in [1, 2)$ we obviously have

$$\|\xi_h\|_p \leq (\mu)^{\frac{d(2-p)}{2p}} \|\xi_h\|_2$$

and, therefore, we can use the upper function found for $p = 2$ for any $p \in [1, 2)$.

Let $\mathbb{B}_{q,r}^s$, $s > 0$, $1 \leq q, r \leq \infty$, denote the Besov space on \mathbb{R}^d , see e.g. Edmunds and Triebel (1996), and later on $\mathbb{H}_q(s, L)$ denote the the ball of the radius $L > 0$ in $\mathbb{B}_{q,\infty}^s$.

Suppose that $K \in \mathbb{H}_\infty(\gamma, L)$ and without loss of generality we assume that $L = 1$ that implies in particular that $\|K\|_\infty \leq 1$.

Theorem 1. Assume that $\gamma > d/2$. Then for any $2 \leq p < \infty$, $h^{(\min)}, h^{(\max)} \in (0, 1)$, and $q \geq 1$

$$\begin{aligned} \mathbb{P} \left\{ \sup_{h \in \mathcal{H}} [\|\xi_h\|_p - C_1 h^{-d/2}] \geq 0 \right\} &\leq C_2 \exp \left\{ -2^{-3/2} \left(h^{(\max)} \right)^{-2d/p} \right\} \\ \mathbb{E} \left\{ \sup_{h \in \mathcal{H}} [\|\xi_h\|_p - C_1 h^{-d/2}] \right\}_+^q &\leq C_3(q) \left(h^{(\max)} \right)^{\frac{qd(2-p)}{2p}} \exp \left\{ -2^{-3/2} \left(h^{(\max)} \right)^{-2d/p} \right\}. \end{aligned}$$

The proof of the theorem is given in Section 5. The constant C_1 involved in the description of found upper functions is bounded function of p on any bounded interval. Thus, the upper functions are independent of p if $p \in [2, p_0]$ for any given $p_0 \geq 2$.

Also it is important to mention that the obtained upper functions are sharp. Indeed, it is not difficult to prove that for any $h > 0$

$$C_4 h^{-d/2} \leq \mathbb{E} \|\xi_h\|_p \leq C_5 h^{-d/2}.$$

This, together with the concentration inequality for gaussian processes, Talagrand (1994), yields in particular for any given $h > 0$

$$\mathbb{P} \left\{ \|\xi_h\|_p \geq \tilde{C}_1 h^{-d/2} \right\} \leq \tilde{C}_2 \exp \left\{ -\tilde{C}_3 h^{-2d/p} \right\}.$$

This inequality coincides, up to numerical constants, with the first inequality in Theorem 1 in particular case when $h^{(\min)} = h^{(\max)} = h$.

3.2. Upper functions for local modulus of continuity under doubling condition

Let metric space (\mathfrak{T}, d) be equipped with Borel measure \varkappa . This measure is doubling if $\exists Q \geq 1$ such that

$$\varkappa\{\mathbb{B}_d(\mathfrak{t}, 2r)\} \leq Q\varkappa\{\mathbb{B}_d(\mathfrak{t}, r)\}, \quad \forall \mathfrak{t} \in \mathfrak{T}, \quad \forall r > 0,$$

where $\mathbb{B}_d(\mathfrak{t}, r)$ is the closed ball with center \mathfrak{t} and radius r . For example, if $\mathfrak{T} = \mathbb{R}^d$ and \varkappa is Lebesgue measure then $Q = 2^d$.

As it was proved in Coifman and Weiss (1971) the existence of a doubling measure on \mathfrak{T} implies that the space \mathfrak{T} is doubling. It means that there exists $N_d \in \mathbb{N}^*$ depending only on Q such that for any $r > 0$ each closed ball in \mathfrak{T} of radius r can be covered by at most N_d closed balls of radius $r/2$. This yields that $\mathbb{B}(\mathfrak{t}, r)$ is totally bounded for any $\mathfrak{t} \in \mathfrak{T}$, $r > 0$ and, moreover,

$$\mathfrak{E}_{\mathbb{B}_d(\mathfrak{t}, r), d}(\delta) \leq \ln(N_d) [\log_2(2r/\delta)]_+, \quad \forall \delta > 0.$$

Let $\mathfrak{t} \in \mathfrak{T}$ and $r > 0$ be fixed. In this section, using Proposition 4 we establish the upper function for local modulus of continuity of χ_θ on $\Theta := \mathbb{B}_d(\mathfrak{t}, r)$. The simplest consequence of this result will be the law of iterated logarithm (LIL) for $|\chi_\theta - \chi_{\mathfrak{t}}|$ as well as its non-asymptotical version. Studying the local modulus of continuity we are obviously interested in the case when r is small even $r \rightarrow 0$. Thus, without loss of generality we will assume that $r \leq 1$.

To apply Proposition 4 we need to define the function g_A , $A(\cdot) = \sqrt{2}\rho(\cdot - \mathfrak{t})$, to choose the function $s_1 \in \mathcal{S}_{a,0}$, and to compute the function $\hat{\mathcal{E}}^{(\varepsilon)}(\Delta, \mathfrak{t})$, $\Delta \in (0, r]$, $\varepsilon \in (0, \sqrt{2} - 1]$, given by

$$\hat{\mathcal{E}}^{(\varepsilon)}(\Delta, \mathfrak{t}) = e_{s_1}^{(a)}\left(g_A([1 + \varepsilon]\Delta), \mathbb{B}_d(\mathfrak{t}, [1 + \varepsilon]\Delta)\right), \quad a = \sqrt{2}\rho.$$

Introduce the function

$$\psi(x) = \sqrt{2} \sup_{\substack{\theta_1, \theta_2 \in \mathbb{B}_d(\mathfrak{t}, 1): \\ d(\theta_1, \theta_2) \leq x}} \rho(\theta_1, \theta_2), \quad x \in (0, 2],$$

and suppose that $\psi(2) < \infty$. Note that obviously $\psi(0) = 0$, since d is a metric, and ψ is increasing. Moreover, for any $u \in (0, r]$

$$g_A^*(u) := \sqrt{2} \sup_{\theta: d(\theta, \mathfrak{t}) \leq u} \rho(\theta, \mathfrak{t}) \leq \psi(u),$$

that allows us to put $g_A = \psi$. Denoting ψ^{-1} the inverse function of ψ we assert that $\forall u \in (0, r]$

$$\mathfrak{E}_{\mathbb{B}_d(\mathfrak{t}, u), \sqrt{2}\rho}(\delta) \leq \mathfrak{E}_{\mathbb{B}_d(\mathfrak{t}, u), d}(\psi^{-1}(\delta)) \leq \ln(N_d) \left[\log_2 \left\{ 2u/\psi^{-1}(\delta) \right\} \right]_+, \quad \forall \delta > 0. \quad (3.1)$$

Hence, if the function ψ is such that Assumption 3 is fulfilled then Proposition 4 is applicable and that provides us with the upper function for $|\chi_\theta - \chi_{\mathfrak{t}}|$ on $\mathbb{B}_d(\mathfrak{t}, r)$. However, this upper function does not admit an explicit expression, in particular its dependence on the variable Δ cannot be analyzed in general. So, we prefer to impose an additional assumption on the function ψ that allows us to obtain the explicit expression of the upper function for $|\chi_\theta - \chi_{\mathfrak{t}}|$ and analyze it as well as the corresponding probability bound for small values of the radius r . We will not be tending here to the maximal generality and suppose that there exist $0 < \underline{c} \leq \bar{c} < \infty$ and $\beta > 0$ such that

$$\underline{c}u^\beta \leq \psi(u) \leq \bar{c}u^\beta, \quad \forall u \in (0, 1]. \quad (3.2)$$

For example, if $d = \rho$ one has $\psi(u) = \sqrt{2}u$, and, therefore, (3.2) holds. Under (3.2) obviously $(\delta/\bar{c})^{\frac{1}{\beta}} \leq \psi^{-1}(\delta) \leq (\delta/\underline{c})^{\frac{1}{\beta}}$ and we get from (3.1)

$$\mathfrak{E}_{\mathbb{B}_d(\mathfrak{t}, u), \sqrt{2}\rho}(\delta) \leq \ln(N_d) \left[\log_2 \left\{ 2u(\delta/\bar{c})^{-\frac{1}{\beta}} \right\} \right]_+, \quad \forall \delta > 0. \quad (3.3)$$

Taking into account that $g_A(u) = \psi(u) \geq \underline{c}u^\beta$ and choosing $s(x) = s^*(x) = (6/\pi^2)(1 + [\ln x]^2)^{-1}$, we obtain from (3.3) for any $\Delta \in (0, r]$

$$\begin{aligned} \widehat{\mathcal{E}}^{(\varepsilon)}(\Delta, \mathfrak{t}) &= \sup_{\delta > 0} \delta^{-2} \mathfrak{E}_{\mathbb{B}_d(\mathfrak{t}, [1+\varepsilon]\Delta), \sqrt{2}\rho} \left(g_A([1+\varepsilon]\Delta) (48\delta)^{-1} s(\delta) \right) \\ &\leq \sup_{\delta > 0} \delta^{-2} \mathfrak{E}_{\mathbb{B}_d(\mathfrak{t}, [1+\varepsilon]\Delta), \sqrt{2}\rho} \left(\underline{c}([1+\varepsilon]\Delta)^\beta (48\delta)^{-1} s(\delta) \right) \\ &\leq \beta^{-1} \ln(N_d) \sup_{\delta > 0} \delta^{-2} \left[\log_2 \left(\frac{96\bar{c}}{\underline{c}} \right) + \log_2 \left(\frac{\delta}{s^*(\delta)} \right) \right]_+ =: C(\beta, \underline{c}, \bar{c}, d). \end{aligned}$$

As we see $\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta, \mathfrak{t})$ is independent of Δ, \mathfrak{t} and bounded from above by the constant which is completely determined by the triplet $(\mathfrak{T}, d, \varkappa)$ and by the quantities \underline{c}, \bar{c} and β . Thus, in view of Proposition 4 the upper function for has the following form (see also Remark 5).

$$\widehat{V}_{s^*}^{(z, \varepsilon)}(\Delta, \mathfrak{t}) = \bar{c}(1 + \varepsilon)^{1+2\beta} \Delta^\beta \sqrt{2[1 + \varepsilon^{-1}]^2 C(\beta, \underline{c}, \bar{c}, d) + [\ell(2(1 + \varepsilon)r/\Delta) + z]},$$

where, remind, $\ell(y) = \ln \{1 + \ln(y)\} + 2 \ln \{1 + \ln \{1 + \ln(y)\}\}$, $y > 0$.

Choose $z = z(r) = \ln \{1 + \ln \{1 + |\ln(r)|\}\}$ and $\varepsilon = \varepsilon(r) := z^{-1}(r)$ and define

$$\begin{aligned} \mathfrak{a}(r) &= (1 + \varepsilon(r))^{1+2\beta} \left(\sup_{\Delta \in (0, r]} \sqrt{\frac{2[1 + \varepsilon^{-1}(r)]^2 C(\beta, \underline{c}, \bar{c}, d) + [\ell(2[1 + \varepsilon(r)]r/\Delta) + z(r)]}{\ln \{1 + |\ln(\Delta)|\}}} \right); \\ \mathfrak{p}(r) &= \frac{2 + [\ln \{1 + \ln(1 + \varepsilon(r))\}]^{-2}}{1 + \ln \{1 + |\ln(r)|\}}. \end{aligned}$$

We note that if $r \rightarrow 0$ then

$$\mathfrak{a}(r) = 1 + \mathcal{O} \left(\frac{\ln \{1 + \ln \{1 + |\ln(r)|\}\}}{\ln \{1 + |\ln(r)|\}} \right), \quad \mathfrak{p}(r) = \mathcal{O} \left(\frac{[\ln \{1 + \ln \{1 + |\ln(r)|\}\}]^2}{\ln \{1 + |\ln(r)|\}} \right). \quad (3.4)$$

The following result is immediate consequence of Proposition 4. Put $\mathfrak{m}(\Delta) = \sup_{\theta \in B_d(\mathfrak{t}, \Delta)} |\chi_\theta - \chi_{\mathfrak{t}}|$.

Theorem 2. *Let \mathfrak{T} be doubling space and suppose that (3.2) holds. Then, we have for any $\mathfrak{t} \in \mathfrak{T}$ and any $r \in (0, 1)$*

$$\mathbb{P} \left\{ \sup_{\Delta \in (0, r]} \left[\frac{\mathfrak{m}(\Delta)}{\bar{c}\Delta^\beta \sqrt{\ln \{1 + |\ln(\Delta)|\}}} \right] \geq \mathfrak{a}(r) \right\} \leq 8\mathfrak{p}(r).$$

The first consequence of Theorem 2 is the law of iterated logarithm. Indeed, taking into account that $\mathfrak{p}(r) \rightarrow 0$, $\mathfrak{a}(r) \rightarrow 1$, $r \rightarrow 0$, we come to the following assertion.

Corollary 2. *Let \mathfrak{T} be doubling space and suppose that (3.2) holds. Then for any $\mathbf{t} \in \mathfrak{T}$ \mathbb{P} -a.s.*

$$\limsup_{\Delta \rightarrow 0+} \left[\frac{\mathbf{m}(\Delta)}{\Delta^\beta \sqrt{\ln \{1 + |\ln(\Delta)|\}}} \right] \leq \bar{c}.$$

We note, that although the statement of Corollary 2 is traditional in probability theory, the non-asymptotical statement of Theorem 2 is much more informative.

The next consequence of Theorem 2 seems more curious. We remark that if \mathfrak{T} is doubling with respect to the intrinsic semi-metric ρ , then the normalizing sequence appeared in the theorem is independent of ρ . Moreover, the function $\mathbf{a}(\cdot)$ depends only on ρ .

Indeed, if $d = \rho$ then $\phi(u) = \sqrt{2}u$, and therefore, $\beta = 1$ and $\underline{c} = \bar{c} = \sqrt{2}$. It yields, in particular, that $C(\beta, \underline{c}, \bar{c}, \rho) = C_{N_\rho}$, where

$$C_{N_\rho} = \ln(N_\rho) \sup_{\delta > 0} \delta^{-2} \left(\left[3 + 2 \log_2(\pi) + \log_2 \left\{ \delta(1 + [\ln \delta]^2) \right\} \right]_+ + 1 \right),$$

and, therefore, $\mathbf{a}(\cdot) = \mathbf{a}_{N_\rho}(\cdot)$, where $\mathbf{a}_{N_\rho}(\cdot)$ is completely determined by ρ via the quantity N_ρ .

Corollary 3. *Let \mathfrak{T} be doubling space with respect to $d = \rho$. Then, for any $\mathbf{t} \in \mathfrak{T}$ and $r \in (0, 1)$*

$$\mathbb{P} \left\{ \sup_{\Delta \in (0, r]} \left[\frac{\mathbf{m}(\Delta)}{\Delta \sqrt{2 \ln \{1 + |\ln(\Delta)|\}}} \right] \geq \mathbf{a}_{N_\rho}(r) \right\} \leq 8\mathbf{p}(r).$$

We note that if \mathfrak{R}^* be the set of metrics ρ for which \mathfrak{T} is doubling and such that $N_\rho \leq N^*$ for some fixed $N^* \in \mathbb{N}^*$ then the function $\mathbf{a}_{N_\rho}(\cdot)$ in the assertion of Corollary 3 can be replaced by the universal on \mathfrak{R}^* function $\mathbf{a}_{N^*}(\cdot)$. The corresponding inequality becomes "metric free".

We finish this section by the consideration of several examples, where the condition (3.2) is verified. In these examples \mathfrak{T} is either \mathbb{R}^d or $\mathbb{R}_+^d, d \geq 1$, ν is Lebesgue measure and d is the euclidian distance.

Example 1. Lévy function Here $\chi_{\mathbf{t}}, \mathbf{t} \in \mathbb{R}^d$, is zero-mean gaussian random field with $\rho = \sqrt{d}$. Hence, (3.2) holds with $\underline{c} = \bar{c} = \sqrt{2}$ and $\beta = 1/2$. We deduce from Theorem 2 that

$$\mathbb{P} \left\{ \sup_{\Delta \in (0, r]} \left[\frac{\mathbf{m}(\Delta)}{\sqrt{2\Delta \ln \{1 + |\ln(\Delta)|\}}} \right] \geq \mathbf{a}(r) \right\} \leq 8\mathbf{p}(r), \quad r \in (0, 1).$$

Example 2. Fractional brownian motion Here $\chi_{\mathbf{t}}, \mathbf{t} \in \mathbb{R}_+$, is zero-mean gaussian random process with $\rho = d^{\alpha/2}, \alpha \in (0, 2]$. Hence, (3.2) holds with $\underline{c} = \bar{c} = \sqrt{2}$ and $\beta = \alpha/2$. We get from Theorem 2

$$\mathbb{P} \left\{ \sup_{\Delta \in (0, r]} \left[\frac{\mathbf{m}(\Delta)}{\sqrt{2\Delta^\alpha \ln \{1 + |\ln(\Delta)|\}}} \right] \geq \mathbf{a}(r) \right\} \leq 8\mathbf{p}(r), \quad r \in (0, 1).$$

Example 3. Ornstein-Uhlenbeck process Here χ_t , $t \in \mathbb{R}_+$, is given by

$$\chi_t = \frac{\sigma}{\sqrt{2\lambda}} b(e^{2\theta t}) e^{-\lambda t}, \quad \lambda, \sigma > 0,$$

where b is the standard Wiener process. In this case $\rho = \lambda^{-1/2} \sigma \sqrt{1 - \exp\{-\lambda d\}}$. Since we consider $r \in (0, 1)$ then (3.2) holds with $\underline{c} = \sigma\sqrt{2}e^{-1}e^{-\lambda/2}$, $\bar{c} = \sigma\sqrt{2}$, $\beta = 1/2$ and we have

$$P \left\{ \sup_{\Delta \in (0, r]} \left[\frac{m(\Delta)}{\sigma \sqrt{2\Delta \ln \{1 + |\ln(\Delta)|\}}} \right] \geq \mathfrak{a}(r) \right\} \leq 8\mathfrak{p}(r), \quad r \in (0, 1).$$

4. Proof of Propositions 1-3

We start this section with establishing an auxiliary result. Let \mathfrak{L} be a set, d_1 and d_2 be semi-metrics on \mathfrak{L} and let \mathcal{L} be a totally bounded subset of \mathfrak{L} with respect to d_1 and d_2 simultaneously. Let $N_i(\delta)$, $\delta > 0$, denote the minimal number of balls of the radius δ in the metric d_i , $i = 1, 2$ needed to cover the set \mathcal{L} .

Lemma 2. *Let $l \in \mathbb{N}^*$ and $\delta_{1,j}, \delta_{2,j} > 0$, $j = \overline{1, l}$ be an arbitrary numbers. One can construct the finite subset $L([\delta_{1,j}, \delta_{2,j}], j = \overline{1, l}) := \{\ell_1, \dots, \ell_N\} \subset \mathcal{L}$ with $N \leq \prod_{j=1}^l N_1(\delta_{1,j}/2) N_2(\delta_{2,j}/2)$ and such that*

$$\forall \ell \in \mathcal{L} \quad \exists \tilde{\ell} \in L([\delta_{1,j}, \delta_{2,j}], j = \overline{1, l}) : \quad d_1(\ell, \tilde{\ell}) \leq \min_{j=\overline{1, l}} \delta_{1,j}, \quad d_2(\ell, \tilde{\ell}) \leq \min_{j=\overline{1, l}} \delta_{2,j}.$$

Proof of Lemma 2 Set $N_{i,j} = N_i(\delta_{i,j}/2)$, $i = 1, 2$, $j = \overline{1, l}$. Since \mathcal{L} is totally bounded in d_i , $i = 1, 2$ there exist $L^{(i,j)} = \{\ell_1^{(i,j)}, \dots, \ell_{N_{i,j}}^{(i,j)}\} \subset \mathcal{L}$ such that

$$\forall \ell \in \mathcal{L} \quad \exists \tilde{\ell} \in L^{(i,j)} : \quad d_i(\ell, \tilde{\ell}) \leq 2^{-1} \delta_{i,j}, \quad i = 1, 2, \quad j = \overline{1, l}.$$

For any $k_i = 1, \dots, N_{i,j}$, $i = 1, 2$, put $\mathcal{L}_{k_i}^{(i,j)} = \{\ell \in \mathcal{L} : d_i(\ell, \ell_{k_i}^{(i,j)}) \leq 2^{-1} \delta_{i,j}\}$ and let $\mathcal{L}_{k_1, k_2}^{(j)} = \mathcal{L}_{k_1}^{(1,j)} \cap \mathcal{L}_{k_2}^{(2,j)}$. First we note that for any $j = \overline{1, l}$

$$\mathcal{L} = \bigcup_{k_1=1}^{N_{1,j}} \bigcup_{k_2=1}^{N_{2,j}} \mathcal{L}_{k_1, k_2}^{(j)} \quad (4.1)$$

Moreover, the construction of $\mathcal{L}_{k_1}^{(1,j)}$ and $\mathcal{L}_{k_2}^{(2,j)}$ implies that

$$d_i(l_1, l_2) \leq \delta_{i,j}, \quad i = 1, 2, \quad \forall l_1, l_2 \in \mathcal{L}_{k_1, k_2}^{(j)}. \quad (4.2)$$

Put $\mathcal{N} = \otimes_{j=1}^l [\{1, \dots, N_{1,j}\} \times \{1, \dots, N_{2,j}\}]$ and define for any $(k^{(1)}, \dots, k^{(l)}) \in \mathcal{N}$

$$\mathcal{L}_{k^{(1)}, \dots, k^{(l)}} = \bigcap_{j=1}^l \mathcal{L}_{k^{(j)}}^{(j)}.$$

The choice of an arbitrary point in each $\mathcal{L}_{k^{(1)}, \dots, k^{(l)}}$ leads to the construction of $L([\delta_1^{(j)}, \delta_2^{(j)}], j = \overline{1, l})$ in view of (4.1) and (4.2).

It remains to note that the cardinality of \mathcal{N} is equal to $\prod_{j=1}^l N_1(\delta_{1,j}/2) N_2(\delta_{2,j}/2)$. ■

4.1. Proof of Proposition 1

I. Probability bound Fix $\vec{s} \in \mathcal{S}_{a,b}$ and put $s_{1,k} = s_1(2^{k/2})$ and $s_{2,k} = s_2(2^k)$, $k \geq 0$. For any $k \geq 0$ put $\delta_1(k) = (24)^{-1} \tilde{\varkappa}_1 2^{-k/2} s_{1,k}$, $\delta_2(k) = (24)^{-1} \tilde{\varkappa}_2 2^{-k} s_{2,k}$ and note that $\delta_1(k), \delta_2(k) \rightarrow 0$, $k \rightarrow \infty$ since $s_1, s_2 \in \mathbb{S}$.

Let $Z_k = L([\delta_1(k), \delta_2(k)])$, $k \geq 0$, be the set constructed in Lemma 2 with $d_1 = a$, $d_2 = b$, $\mathcal{L} = \tilde{\Theta}$ and $l = 1$. By N_k , $k \geq 0$, we denote the cardinality of Z_k .

Fix $\varepsilon > 0$ and put $\epsilon = \varepsilon/(1 + \varepsilon)$, $k_0 = \lfloor 2 \ln_2(1/\epsilon) \rfloor + 1$. Let θ_m , $m = 1, \dots, N_{k_0}$, be the elements of Z_{k_0} . For any $m = 1, \dots, N_{k_0}$ define

$$\Theta^{(m)} = \left\{ \theta \in \tilde{\Theta} : a(\theta, \theta_m) \leq \delta_1(k_0), b(\theta, \theta_m) \leq \delta_2(k_0) \right\},$$

and remark that the definition of the sets Z_{k_0} implies that $\tilde{\Theta} = \bigcup_{m=1}^{N_{k_0}} \Theta^{(m)}$.

In view of the last remark we get

$$\mathbb{P} \left\{ \sup_{\theta \in \tilde{\Theta}} \Psi(\chi_\theta) \geq U_{\vec{s}}^{(\varepsilon)}(y, \tilde{\varkappa}, \tilde{\Theta}) \right\} \leq \sum_{m=1}^{N_{k_0}} \mathbb{P} \left\{ \sup_{\theta \in \Theta^{(m)}} \Psi(\chi_\theta) \geq U_{\vec{s}}^{(\varepsilon)}(y, \tilde{\varkappa}, \tilde{\Theta}) \right\}. \quad (4.3)$$

For any $\theta \in \tilde{\Theta}$ let $z_k(\theta)$ be an arbitrary element of Z_k satisfying

$$a(\theta, z_k(\theta)) \leq \delta_1(k), \quad b(\theta, z_k(\theta)) \leq \delta_2(k). \quad (4.4)$$

Fix $m = 1, \dots, N_{k_0}$. The continuity of the mapping $\theta \mapsto \chi_\theta$ guarantees that P-a.s.

$$\chi_\theta = \chi_{\theta_m} + \sum_{k=k_0+1}^{\infty} [\chi_{z_k(\theta)} - \chi_{z_{k-1}(\theta)}], \quad \forall \theta \in \Theta^{(m)}, \quad (4.5)$$

where $z_{k_0}(\theta) = \theta_m$, $\forall \theta \in \Theta^{(m)}$. Note also that independently of θ for all $k \geq k_0 + 1$

$$a(z_k(\theta), z_{k-1}(\theta)) \leq \delta_1(k) + \delta_1(k-1) =: \tilde{\delta}_1(k), \quad (4.6)$$

$$b(z_k(\theta), z_{k-1}(\theta)) \leq \delta_2(k) + \delta_2(k-1) =: \tilde{\delta}_2(k). \quad (4.7)$$

This is the simplest consequence of triangle inequality and (4.4). Introduce the sequence $c_k, k \geq 1$:

$$c_k = 4^{-1} \max \{s_{1,k}, s_{1,k-1}, s_{2,k}, s_{2,k-1}\}$$

and remark that $\sum_{k \geq 1} c_k \leq 1$ that follows from the assumption $s_1, s_2 \in \mathbb{S}$.

We get from sub-additivity of Ψ , (4.5), (4.6) and (4.7) for any $\theta \in \Theta^{(m)}$

$$\Psi(\chi_\theta) \leq \Psi(\chi_{\theta_m}) + \sup_{k \geq k_0+1} \sup_{\substack{(u,v) \in Z_k \times Z_{k-1}: \\ a(u,v) \leq \tilde{\delta}_1(k), b(u,v) \leq \tilde{\delta}_2(k)}} c_k^{-1} \Psi(\chi_u - \chi_v), \quad (4.8)$$

To simplify the notations we will write U instead of $U_{\vec{s}}^{(\varepsilon)}(y, \tilde{\varkappa}, \tilde{\Theta})$ and \mathcal{E} instead of $e_{\vec{s}}(\tilde{\varkappa}, \tilde{\Theta})$.

We obtain from (4.8)

$$\begin{aligned} \mathbb{P} \left\{ \sup_{\theta \in \Theta^{(m)}} \Psi(\chi_\theta) \geq U \right\} &\leq \mathbb{P} \left\{ \Psi(\chi_{\theta_m}) \geq U(1 + \varepsilon)^{-1} \right\} \\ &+ \sum_{k=k_0+1}^{\infty} \sum_{\substack{(u,v) \in Z_k \times Z_{k-1}: \\ a(u,v) \leq \tilde{\delta}_1(k), b(u,v) \leq \tilde{\delta}_2(k)}} \mathbb{P} \left\{ \Psi(\chi_u - \chi_v) \geq \epsilon U c_k \right\} =: I_1 + I_2. \end{aligned} \quad (4.9)$$

We have in view of Assumption 1 (1)

$$\begin{aligned}
\mathbb{P} \left\{ \Psi(\chi_{\theta_m}) \geq U(1+\varepsilon)^{-1} \right\} &\leq \mathbb{c} \exp \left\{ -\frac{(1+\varepsilon)^{-2}U^2}{\{A(\theta_m)\}^2 + (1+\varepsilon)^{-1}UB(\theta_m)} \right\} \\
&\leq \mathbb{c} \exp \left\{ -\frac{(1+\varepsilon)^{-2}U^2}{\tilde{\varkappa}_1^2 + U\tilde{\varkappa}_2} \right\} \leq \mathbb{c} \exp \left\{ -(1+\varepsilon)^{-2}(y + 2\epsilon^{-2}\mathcal{E}) \right\} \\
&\leq \mathbb{c} \exp \left\{ -(1+\varepsilon)^{-2}y - \epsilon^{-2}\mathcal{E} \right\}.
\end{aligned} \tag{4.10}$$

In order to get (4.10) we have first used that $\tilde{\varkappa}_1 \geq \sup_{\theta \in \tilde{\Theta}} A(\theta)$, $\tilde{\varkappa}_2 \geq \sup_{\theta \in \tilde{\Theta}} B(\theta)$. Next, we have used that $U \geq v$, where v is the maximal root of the equation

$$\frac{u^2}{\tilde{\varkappa}_1^2 + u\tilde{\varkappa}_2} = y + 2\epsilon^{-2}\mathcal{E}. \tag{4.11}$$

We also have used that $(1+\varepsilon)^{-2} \geq 1/2$.

In view of Assumptions 1 (2), 2, (4.6) and (4.7) for any $u, v \in Z_k \times Z_{k-1}$ satisfying $a(u, v) \leq \tilde{\delta}_1(k)$, $b(u, v) \leq \tilde{\delta}_2(k)$ we have

$$\begin{aligned}
\mathbb{P} \left\{ \Psi(\chi_u - \chi_v) \geq \epsilon U c_k \right\} &\leq \mathbb{c} \exp \left\{ -\frac{(\epsilon U c_k)^2}{\{a(u, v)\}^2 + (\epsilon U c_k) b(u, v)} \right\} \\
&\leq \mathbb{c} \exp \left\{ -\frac{(\epsilon U)^2}{\{\tilde{\delta}_1(k) c_k^{-1}\}^2 + U \{\tilde{\delta}_2(k) c_k^{-1}\}} \right\}.
\end{aligned} \tag{4.12}$$

Here we have used that $\epsilon < 1$. Let us remark that

$$\begin{aligned}
\tilde{\delta}_1(k) c_k^{-1} &\leq 4(24)^{-1} \tilde{\varkappa}_1 \left(2^{-k/2} s_{1,k} + 2^{-(k-1)/2} s_{1,k-1} \right) \min \left\{ s_{1,k}^{-1}, s_{1,k-1}^{-1} \right\} \leq \tilde{\varkappa}_1 2^{-k/2-1}; \\
\tilde{\delta}_2(k) c_k^{-1} &\leq 4(24)^{-1} \tilde{\varkappa}_2 \left(2^{-k} s_{2,k} + 2^{-k+1} s_{2,k-1} \right) \min \left\{ s_{2,k}^{-1}, s_{2,k-1}^{-1} \right\} \leq \tilde{\varkappa}_2 2^{-k-1}.
\end{aligned}$$

Thus, continuing (4.12) we obtain

$$\mathbb{P} \left\{ \Psi(\chi_u - \chi_v) \geq \epsilon U c_k \right\} \leq \mathbb{c} \exp \left\{ -\frac{2^{k+1} \epsilon^2 U^2}{\tilde{\varkappa}_1^2 + U \tilde{\varkappa}_2} \right\} \leq \mathbb{c} \exp \left\{ -2^{k+1} \epsilon^2 (y + 2\epsilon^{-2}\mathcal{E}) \right\}. \tag{4.13}$$

Here we have used (4.11). Noting that the right hand side of (4.13) does not depend on u, v we get

$$\begin{aligned}
I_2 &\leq \mathbb{c} \sum_{k=k_0+1}^{\infty} N_k N_{k-1} \exp \left\{ -2^{k+1} \epsilon^2 (y + 2\epsilon^{-2}\mathcal{E}) \right\} \\
&\leq \mathbb{c} \exp(-y) \sum_{k=k_0+1}^{\infty} N_k N_{k-1} \exp \left\{ -2^{k+2} \mathcal{E} - 2^{k-k_0} \right\}.
\end{aligned} \tag{4.14}$$

Here we have used the definition of k_0 and that $y \geq 1$.

Let us make several remarks. First we note that in view of Lemma 2

$$\ln(N_k) \leq \mathfrak{E}_{\tilde{\Theta}, a} \left((24)^{-1} \tilde{\varkappa}_1 2^{-1-k/2} s_{1,k} \right) + \mathfrak{E}_{\tilde{\Theta}, b} \left((24)^{-1} \tilde{\varkappa}_2 2^{-k-1} s_{2,k} \right).$$

Taking into account that $s_{1,k} = s_1(2^{k/2})$ and denoting $\delta_1 = 2^{k/2}$ we obtain from (1.9)

$$\mathfrak{E}_{\tilde{\Theta}, a} \left((24)^{-1} \tilde{\varkappa}_1 2^{-1-k/2} s_{1,k} \right) = \mathfrak{E}_{\tilde{\Theta}, a} \left(\tilde{\varkappa}_1 (48\delta_1)^{-1} s_1(\delta_1) \right) \leq \delta_1^2 e_{s_1}^{(a)}(\tilde{\varkappa}_1, \tilde{\Theta}) = 2^k e_{s_1}^{(a)}(\tilde{\varkappa}_1, \tilde{\Theta}).$$

Taking into account that $s_{2,k} = s_1(2^k)$ and denoting $\delta_2 = 2^k$ we obtain from (1.9)

$$\mathfrak{E}_{\tilde{\Theta}, b} \left((24)^{-1} \tilde{\varkappa}_2 2^{-1-k} s_{2,k} \right) = \mathfrak{E}_{\tilde{\Theta}, b} \left(\tilde{\varkappa}_2 (48\delta_2)^{-1} s_2(\delta_2) \right) \leq \delta_2 e_{s_2}^{(b)}(\tilde{\varkappa}_2, \tilde{\Theta}) = 2^k e_{s_2}^{(b)}(\tilde{\varkappa}_2, \tilde{\Theta}).$$

Thus, we have for any $k \geq 1$

$$\ln(N_k N_{k-1}) \leq 2^{k+1} \left[e_{s_1}^{(a)}(\tilde{\Theta}) + e_{s_2}^{(b)}(\tilde{\Theta}) \right] = 2^{k+1} \mathcal{E} \quad (4.15)$$

and, therefore, $\forall k \geq k_0 + 1$

$$\ln(N_k N_{k-1}) - 2^{k+2} \mathcal{E} \leq -2^{k_0+2} \mathcal{E} \leq -4\epsilon^{-2} \mathcal{E}.$$

It yields together with (4.14)

$$I_2 \leq c \exp \left\{ -y - 4\epsilon^{-2} \mathcal{E} \right\}. \quad (4.16)$$

We get from (4.9), (4.10) and (4.16) for any $m = 1, \dots, N_{k_0}$

$$\mathbb{P} \left\{ \sup_{\theta \in \Theta^{(m)}} \Psi(\chi_\theta) \geq U \right\} \leq 2c \exp \left\{ -y(1+\epsilon)^{-2} - \epsilon^{-2} \mathcal{E} \right\}. \quad (4.17)$$

The last bound is independent of m and we have from (4.3) and (4.17)

$$\mathbb{P} \left\{ \sup_{\theta \in \tilde{\Theta}} \Psi(\chi_\theta) \geq U \right\} \leq 2c N_{k_0} \exp \left\{ -y(1+\epsilon)^{-2} - \epsilon^{-2} \mathcal{E} \right\}.$$

It remains to note that $(1 - 3\epsilon)^2 = (1 + \epsilon)^{-1}$ and that, similarly to (4.15),

$$\ln(N_{k_0}) \leq 2^{k_0} \mathcal{E} \leq \epsilon^2 \mathcal{E},$$

and we come to the first assertion of the proposition.

II. Moment's bound We get for any $y \geq 1$

$$\begin{aligned} E &:= \mathbb{E} \left(\sup_{\theta \in \tilde{\Theta}} \Psi(\chi_\theta) - U \right)_+^q = q \int_0^\infty x^{q-1} \mathbb{P} \left\{ \sup_{\theta \in \tilde{\Theta}} \Psi(\chi_\theta) \geq U + x \right\} dx \\ &= q [U]^q \int_0^\infty v^{q-1} \mathbb{P} \left\{ \sup_{\theta \in \tilde{\Theta}} \Psi(\chi_\theta) \geq (1+v)U \right\} dv. \end{aligned} \quad (4.18)$$

Note that $(1+v)U \geq U_s^{(\epsilon)}((1+v)y, \tilde{\varkappa}, \tilde{\Theta})$. Therefore, applying the first statement of the proposition, where y is replaced by vy we obtain from (4.18)

$$E \leq 2c\Gamma(q+1) \left[(1+\epsilon)^2 y^{-1} U \right]^q \exp \left\{ -y/(1+\epsilon)^2 \right\}.$$

■

4.2. Proof of Proposition 2

We start with establishing some technical results used in the sequel.

Preliminaries ¹⁰. First we formulate the simple consequence of Proposition 1.

Let Θ_1, Θ_2 be given subsets of Θ . For any $\vec{s} = (s_1, s_2) \in \mathcal{S}_{a,b}$ and any $\varkappa = (\varkappa_1, \varkappa_2) \in \mathbb{R}_+^2 \setminus \{0\}$ introduce the following quantity

$$e_{\vec{s}}(\varkappa, \Theta_1, \Theta_2) = e_{s_1}^{(a)}(\varkappa_1, \Theta_1) + e_{s_2}^{(b)}(\varkappa_2, \Theta_2).$$

Put any $\varepsilon > 0$ and any $y \geq 0$

$$U_{\vec{s}}^{(\varepsilon)}(y, \varkappa, \Theta_1, \Theta_2) = \varkappa_1 \sqrt{2[1 + \varepsilon^{-1}]^2 e_{\vec{s}}(\varkappa, \Theta_1, \Theta_2) + y} + \varkappa_2 (2[1 + \varepsilon^{-1}]^2 e_{\vec{s}}(\varkappa, \Theta_1, \Theta_2) + y).$$

Lemma 3. *Let Assumptions 1-3 hold and let Θ_1, Θ_2 be given subsets of Θ . Let \varkappa be chosen such that $\varkappa_1 \geq \sup_{\theta \in \Theta_1} A(\theta)$ and $\varkappa_2 \geq \sup_{\theta \in \Theta_2} B(\theta)$. Then $\forall \vec{s} \in \mathcal{S}_{a,b}$, $\forall \varepsilon \in (0, \sqrt{2} - 1]$ and $\forall y \geq 1$,*

$$\mathbb{P} \left\{ \sup_{\theta \in \Theta_1 \cap \Theta_2} \Psi(\chi_\theta) \geq U_{\vec{s}}^{(\varepsilon)}(y, \varkappa, \Theta_1, \Theta_2) \right\} \leq 2c \exp \left\{ -y/(1 + \varepsilon)^2 \right\}.$$

Moreover, for any $q \geq 1$, putting $C_{\varepsilon,q} = 2c\Gamma(q+1)(1 + \varepsilon)^{2q}$, one has

$$\mathbb{E} \left\{ \sup_{\theta \in \Theta_1 \cap \Theta_2} \Psi(\chi_\theta) - U_{\vec{s}}^{(\varepsilon)}(y, \varkappa, \Theta_1, \Theta_2) \right\}_+^q \leq C_{\varepsilon,q} \left[y^{-1} U_{\vec{s}}^{(\varepsilon)}(y, \varkappa, \Theta_1, \Theta_2) \right]^q \exp \left\{ -\frac{y}{(1 + \varepsilon)^2} \right\}.$$

To prove the lemma it suffices to note the following simple facts. In view of the assumptions imposed on \varkappa and obvious inclusions $\Theta_1 \cap \Theta_2 \subseteq \Theta_1$, $\Theta_1 \cap \Theta_2 \subseteq \Theta_2$ we have

$$\varkappa_1 \geq \sup_{\theta \in \Theta_1 \cap \Theta_2} A(\theta), \quad \varkappa_2 \geq \sup_{\theta \in \Theta_1 \cap \Theta_2} B(\theta), \quad e_{\vec{s}}(\varkappa, \Theta_1, \Theta_2) \geq e_{\vec{s}}(\varkappa, \Theta_1 \cap \Theta_2).$$

It yields $U_{\vec{s}}^{(\varepsilon)}(y, \varkappa, \Theta_1, \Theta_2) \geq U_{\vec{s}}^{(\varepsilon)}(y, \varkappa, \Theta_1 \cap \Theta_2)$ and to get the assertion of the lemma we apply Proposition 1 with $\tilde{\Theta} = \Theta_1 \cap \Theta_2$.

²⁰. Note that $\Psi(\chi_\bullet) : \Theta \rightarrow \mathbb{R}_+$ is obviously P-a.s. continuous in $\mathfrak{a} \vee \mathfrak{b}$ as a composition of two continuous mappings between metric spaces. Hence Corollary 1 is applicable with $\mathfrak{T} = \Theta$, $\mathfrak{d} = \mathfrak{a} \vee \mathfrak{b}$, $(\Omega, \mathfrak{B}, \mathbf{P}) = (\Omega, \mathfrak{B}, \mathbf{P})$ $\zeta(\mathfrak{t}, \cdot) = \Psi(\chi_{\theta}(\cdot))$ and $g(\mathfrak{t})$ is either $V_{\vec{s}}^{(z,\varepsilon)}(\theta)$ or $U_{\vec{s}}^{(z,\varepsilon,q)}(\theta)$, $\mathfrak{t} = \theta$.

Proof of the proposition Put $\delta_l = (1 + \varepsilon)^l$, $l \geq 0$, and introduce the following sets

$$\Theta_A^{(l)} = \{\theta \in \Theta : \underline{A}\delta_{l-1} \leq A(\theta) \leq \underline{A}\delta_l\}, \quad \Theta_B^{(l)} = \{\theta \in \Theta : \underline{B}\delta_{l-1} \leq B(\theta) \leq \underline{B}\delta_l\}, \quad l \in \mathbb{N}^*.$$

The idea is to apply Lemma 3 with $\Theta_1 = \Theta_A^{(j)}$ and $\Theta_2 = \Theta_B^{(k)}$ for any given $j, k \geq 1$. To do that we will need to bound from below $V_{\vec{s}}^{(z,\varepsilon)}(\theta)$ and $U_{\vec{s}}^{(z,\varepsilon,q)}(\theta)$ on $\Theta_A^{(j)} \cap \Theta_B^{(k)}$. We will consider only j, k such that $\Theta_A^{(j)} \cap \Theta_B^{(k)} \neq \emptyset$ and supremum over empty set is assumed to be zero. Also we will accept the following agreement: if $B \equiv 0$, $b \equiv 0$ then $k \equiv 0$ and $\Theta_B^{(0)} = \hat{\Theta}$.

Probability bound Let $\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}$ be fixed and put $u = \mathcal{A}_\varepsilon(\theta)$, $v = \mathcal{B}_\varepsilon(\theta)$. Note that

$$e_{s_1}^{(a)}(\underline{A}u, \Theta_A(\underline{A}u)) \geq e_{s_1}^{(a)}(\underline{A}u, \Theta_A(\underline{A}\delta_j)) \geq e_{s_1}^{(a)}(\underline{A}u, \Theta_A^{(j)}) \geq e_{s_1}^{(a)}(\underline{A}\delta_{j+1}, \Theta_A^{(j)}). \quad (4.19)$$

To get the first two inequalities in (4.19) we have used that $\Theta_A^{(j)} \subseteq \Theta_A(\underline{A}\delta_j) \subseteq \Theta_A(\underline{A}u)$ in view of $\delta_j \leq u$ since $\theta \in \Theta_A^{(j)}$. To get the last inequality in (4.19) we have used that the entropy is decreasing function of its argument and that $\delta_{j+1} \geq u$ since $\theta \in \Theta_A^{(j)}$.

By the same reasons

$$e_{s_2}^{(b)}(\underline{B}v, \Theta_B(\underline{B}v)) \geq e_{s_2}^{(b)}(\underline{B}v, \Theta_B(\underline{B}\delta_k)) \geq e_{s_2}^{(b)}(\underline{B}v, \Theta_B^{(k)}) \geq e_{s_2}^{(b)}(\underline{B}\delta_{k+1}, \Theta_B^{(k)}). \quad (4.20)$$

Taking into account that left hand sides in (4.19) and (4.20) are independent of θ , whenever $\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}$, we deduce from (2.3), (4.19) and (4.20)

$$\begin{aligned} \mathcal{E}_\varepsilon(\mathcal{A}_\varepsilon(\theta), \mathcal{B}_\varepsilon(\theta)) &:= e_{s_1}^{(a)}(\underline{A}u, \Theta_A(\underline{A}u)) + e_{s_2}^{(b)}(\underline{B}v, \Theta_B(\underline{B}v)) \\ &\geq e_{s_1}^{(a)}(\underline{A}\delta_{j+1}, \Theta_A^{(j)}) + e_{s_2}^{(b)}(\underline{B}\delta_{k+1}, \Theta_B^{(k)}) = e_{\vec{s}}(\varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}), \end{aligned} \quad (4.21)$$

for any $\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}$, where we put $\varkappa = (\underline{A}\delta_{j+1}, \underline{B}\delta_{k+1})$.

We obtain from (4.21), putting $y = (1 + \varepsilon)^2[z + \ell(\delta_j) + \ell(\delta_k)]$,

$$V_{\vec{s}}^{(z, \varepsilon)}(\theta) \geq U_{\vec{s}}^{(\varepsilon)}(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}), \quad \forall \theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}. \quad (4.22)$$

Here we have also used that obviously $(1 + \varepsilon)^2 A(\theta) \geq \underline{A}\delta_{j+1} =: \varkappa_1$ and $(1 + \varepsilon)^2 B(\theta) \geq \underline{B}\delta_{k+1} =: \varkappa_2$ for any $\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}$. Moreover, we have used that $2[1 + \varepsilon^{-1}]^2 \geq 4\varepsilon^{-2}$ for any $\varepsilon \in (0, \sqrt{2} - 1]$.

Therefore, we obtain $\forall j, k \geq 1$ in view of (4.22)

$$\Psi_{j,k}^*(z) := \sup_{\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}} \left\{ \Psi(\chi_\theta) - V_{\vec{s}}^{(z, \varepsilon)}(\theta) \right\} \leq \sup_{\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}} \left\{ \Psi(\chi_\theta) - U_{\vec{s}}^{(\varepsilon)}(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}) \right\}.$$

Since $\varkappa_1 := \underline{A}\delta_{j+1} \geq \sup_{\theta \in \Theta_A^{(j)}} A(\theta)$ and $\varkappa_2 := \underline{B}\delta_{k+1} \geq \sup_{\theta \in \Theta_B^{(k)}} B(\theta)$, Lemma 3 is applicable with $\Theta_1 = \Theta_A^{(j)}$ and $\Theta_2 = \Theta_B^{(k)}$. Thus, applying it we obtain $\forall j, k \geq 1$ and $\forall z \geq 1$

$$\mathbb{P} \left\{ \Psi_{j,k}^*(z) \geq 0 \right\} \leq 2c \exp \{-z\} w_j w_k, \quad (4.23)$$

where we put $w_m = \left[1 + m \ln(1 + \varepsilon) \right]^{-1} \left[1 + \ln \{ 1 + m \ln(1 + \varepsilon) \} \right]^{-2}$. Noting that

$$\sum_{m=1}^{\infty} w_m \leq 1 + \left[\ln \{ 1 + \ln(1 + \varepsilon) \} \right]^{-2},$$

taking into account that the union of $\left\{ \Theta_A^{(j)} \cap \Theta_B^{(k)}, j, k \geq 1, \right\}$ covers Θ and summing up the right hand side in (4.23) over j, k , we arrive at the first assertion of the proposition.

Moment's bound Using the same arguments having led to (4.22) and taking into account (4.21) we obtain for any $\theta \in \Theta^{(j,k)}$ and any $q \geq 1$

$$U_{\vec{s}}^{(z,\varepsilon,q)}(\theta) \geq U_{\vec{s}}^{(\varepsilon)}\left(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\right),$$

where $y = (1 + \varepsilon)^2 \left[z + \tau_\varepsilon e_{\vec{s}}\left(\varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\right) + (\varepsilon + q) \ln(\delta_j \delta_k) \right]$, $\tau_\varepsilon = 2(2 + \varepsilon)\varepsilon^{-1}$.

Thus, applying the second assertion of Lemma 3, $\forall j, k \geq 1$ and $\forall z \geq 1$

$$\begin{aligned} E_{j,k} &:= \mathbb{E} \left(\sup_{\theta \in \Theta^{(j,k)}} \Psi(\chi_\theta) - U_{\vec{s}}^{(\varepsilon)}\left(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\right) \right)_+^q \\ &\leq 2c\Gamma(q+1) \left[\frac{U_{\vec{s}}^{(\varepsilon)}\left(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\right)}{z + \tau_\varepsilon e_{\vec{s}}\left(\varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\right) + (\varepsilon + q) \ln(\delta_j \delta_k)} \right]^q (\delta_j \delta_k)^{-\varepsilon-q} \exp\{-z\}. \end{aligned} \quad (4.24)$$

Putting for brevity $e_{j,k} = e_{\vec{s}}\left(\varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\right)$ and noting that for any $j, k \geq 1$ and $z \geq 1$

$$U_{\vec{s}}^{(\varepsilon)}\left(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\right) \leq 2[\varkappa_1 \vee \varkappa_2] \left[4\varepsilon^{-2} e_{j,k} + y \right] \leq 2[\underline{A} \vee \underline{B}] [\delta_{j+1} \vee \delta_{k+1}] [2(1 + \varepsilon^{-1})^2 e_{j,k} + y],$$

and that for any $\varepsilon \in (0, \sqrt{2} - 1]$

$$\frac{2(1 + \varepsilon^{-1})^2 e_{j,k} + (1 + \varepsilon)^2 \left[z + \tau_\varepsilon e_{j,k} + (\varepsilon + q) \ln(\delta_j \delta_k) \right]}{z + \tau_\varepsilon e_{j,k} + (\varepsilon + q) \ln(\delta_j \vee \delta_k)} \leq (1 + \varepsilon)^4 (2\varepsilon)^{-1},$$

we get finally from (4.24) for $j, k \geq 1$

$$E_{j,k} \leq 2c\Gamma(q+1) [\varepsilon^{-1}(1 + \varepsilon)^5]^q [\underline{A} \vee \underline{B}]^q (\delta_j \delta_k)^{-\varepsilon} \exp\{-z\}. \quad (4.25)$$

Here we have used that $(\delta_j \vee \delta_k)/(\delta_j \delta_k) \leq 1$ since $\delta_j, \delta_k \geq 1$. Taking into account that

$$\mathbb{E} \left(\sup_{\theta \in \Theta} \left\{ \Psi(\chi_\theta) - U_{\vec{s}}^{(z,\varepsilon,q)}(\theta) \right\} \right)_+^q \leq \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} E_{j,k}$$

and that $\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} (\delta_j \delta_k)^{-\varepsilon} = [(1 + \varepsilon)^\varepsilon - 1]^{-2}$, we come to the second assertion of the proposition in view of (4.25), where one can replace $(1 + \varepsilon)^5$ par $2^{5/2}$ since $\varepsilon \in (0, \sqrt{2} - 1]$. We have also used that $[(1 + \varepsilon)^\varepsilon - 1]^{-2} \leq 2\varepsilon^{-4}$ since $\varepsilon \in (0, \sqrt{2} - 1]$. ■

4.3. Proof of Proposition 3

First we discuss the measurability issue. Note that $\Psi(\chi_\bullet) : \Theta \rightarrow \mathbb{R}_+$ is obviously P-a.s. continuous in a $\vee \mathfrak{b}$ as a composition of two continuous mappings between metric spaces. Hence Lemma 1 is applicable with $\mathfrak{T} = \Theta$, $\mathfrak{d} = \mathfrak{a} \vee \mathfrak{b}$, $(\Omega, \mathfrak{B}, \mathbf{P}) = (\Omega, \mathfrak{B}, \mathbf{P})$ $\zeta(\mathfrak{t}, \cdot) = \Psi(\chi_{\theta(\cdot)})$, $\mathfrak{t} = \theta$, $\mathfrak{Z} = \mathfrak{A}$, $\mathfrak{T}_\zeta = \Theta_\alpha$ and $g(\mathfrak{z}) = \widehat{U}^{(z,\varepsilon,r)}(\alpha)$, $\mathfrak{z} = \alpha$.

The proof of the proposition is similar to those of Proposition 2. Let $\varepsilon \in (0, \sqrt{2} - 1]$ be fixed and put $\delta_l = (1 + \varepsilon)^{-l}$, $l \geq 0$, $\delta_l = 1$, $l < 0$. Introduce the following sets:

$$\mathfrak{A}_1^{(l)} = \{\alpha \in \mathfrak{A} : \bar{\tau}_1 \delta_{l+1} \leq \tau_1(\alpha) \leq \bar{\tau}_1 \delta_l\}, \quad \mathfrak{A}_2^{(l)} = \{\alpha \in \mathfrak{A} : \bar{\tau}_2 \delta_{l+1} \leq \tau_2(\alpha) \leq \bar{\tau}_2 \delta_l\}, \quad l \in \mathbb{N}.$$

Fix $j, k \geq 0$, $\alpha \in \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}$ and put $u = (1 + \varepsilon)\tau_1(\alpha)$, $v = (1 + \varepsilon)\tau_2(\alpha)$. Note that

$$\begin{aligned} e_{s_1(u, \cdot)}^{(a)}(\lambda_1^{-1} g_A(u), \Theta'_1(u)) &\geq e_{s_1(u, \cdot)}^{(a)}(\lambda_1^{-1} g_A(u), \Theta'_1(\bar{\tau}_1 \delta_j)) \\ &\geq e_{s_1(u, \cdot)}^{(a)}(\lambda_1^{-1} g_A(\bar{\tau}_1 \delta_{j-1}), \Theta'_1(\bar{\tau}_1 \delta_j)). \end{aligned} \quad (4.26)$$

To get the first inequality in (4.26) we have used that $\Theta'_1(\bar{\tau}_1 \delta_j) \subseteq \Theta'_1(u)$ in view of $\bar{\tau}_1 \delta_j \leq u$ since $\alpha \in \mathfrak{A}_1^{(j)}$. To get the second inequality in (4.26) we have used that $g_A(\bar{\tau}_1 \delta_{j-1}) \geq g_A(u)$, since g_A is increasing and $\bar{\tau}_1 \delta_{j-1} \geq u$ for $\alpha \in \mathfrak{A}_1^{(j)}$. Moreover we have used that the entropy is decreasing function of its argument.

Remembering the definition of $e_{s_1(u, \cdot)}^{(a)}(\lambda_1^{-1} g_A(\bar{\tau}_1 \delta_{j-1}), \Theta'_1(\bar{\tau}_1 \delta_j))$, see (1.9), we have

$$\begin{aligned} e_{s_1(u, \cdot)}^{(a)}(\lambda_1^{-1} g_A(\bar{\tau}_1 \delta_{j-1}), \Theta'_1(\bar{\tau}_1 \delta_j)) &= \sup_{\delta > 0} \delta^{-2} \mathfrak{E}_{\Theta'_1(\bar{\tau}_1 \delta_j), a}(\lambda_1^{-1} g_A(\bar{\tau}_1 \delta_{j-1})(48\delta)^{-1} s_1(u, \delta)) \\ &= \sup_{\delta > 0} \delta^{-2} \mathfrak{E}_{\Theta'_1(\bar{\tau}_1 \delta_j), a} \left(\lambda_1^{-1} g_A(\bar{\tau}_1 \delta_{j-1})(48\delta)^{-1} s_1(\bar{\tau}_1 \delta_j, \delta) \left[\frac{s_1(u, \delta)}{s_1(\bar{\tau}_1 \delta_j, \delta)} \right] \right) \\ &\geq \sup_{\delta > 0} \delta^{-2} \mathfrak{E}_{\Theta'_1(\bar{\tau}_1 \delta_j), a} (g_A(\bar{\tau}_1 \delta_{j-1})(48\delta)^{-1} s_1(\bar{\tau}_1 \delta_j, \delta)) \\ &=: e_{s_1(\bar{\tau}_1 \delta_j, \cdot)}^{(a)}(g_A(\bar{\tau}_1 \delta_{j-1}), \Theta'_1(\bar{\tau}_1 \delta_j)). \end{aligned} \quad (4.27)$$

To get (4.27) we have used that $1 \leq u/\bar{\tau}_1 \delta_j \leq 1 + \varepsilon \leq \sqrt{2}$, the definition of λ_1 and, as previously, that the entropy is decreasing function of its argument. We obtain from (4.26) and (4.27)

$$e_{s_1(u, \cdot)}^{(a)}(\lambda_1^{-1} g_A(u), \Theta'_1(u)) \geq e_{s_1(\bar{\tau}_1 \delta_j, \cdot)}^{(a)}(g_A(\bar{\tau}_1 \delta_{j-1}), \Theta'_1(\bar{\tau}_1 \delta_j)). \quad (4.28)$$

By the same reasons we have

$$e_{s_2(v, \cdot)}^{(b)}(\lambda_2^{-1} g_B(v), \Theta'_2(v)) \geq e_{s_2(\bar{\tau}_2 \delta_k, \cdot)}^{(b)}(g_B(\bar{\tau}_2 \delta_{k-1}), \Theta'_2(\bar{\tau}_2 \delta_k)), \quad (4.29)$$

and, we get from (4.28) and (4.29) for any $\alpha \in \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}$

$$\begin{aligned} \widehat{\mathcal{E}}^{(\varepsilon)}(\alpha) &:= \mathcal{E}'(u, v) \geq e_{s_1(\bar{\tau}_1 \delta_j, \cdot)}^{(a)}(g_A(\bar{\tau}_1 \delta_{j-1}), \Theta'_1(\bar{\tau}_1 \delta_j)) + e_{s_2(\bar{\tau}_2 \delta_k, \cdot)}^{(b)}(g_B(\bar{\tau}_2 \delta_{k-1}), \Theta'_2(\bar{\tau}_2 \delta_k)) \\ &= e_{\vec{s}}(\varkappa, \Theta'_1(\bar{\tau}_1 \delta_j), \Theta'_2(\bar{\tau}_2 \delta_k)), \end{aligned} \quad (4.30)$$

where we have put $\varkappa = (g_A(\bar{\tau}_1 \delta_{j-1}), g_B(\bar{\tau}_2 \delta_{k-1}))$ and $\vec{s} = (s_1(\bar{\tau}_1 \delta_j, \cdot), s_2(\bar{\tau}_2 \delta_k, \cdot))$.

Note also that for any $\alpha \in \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}$ in view of monotonicity of functions g_A and g_B

$$g_A([1 + \varepsilon]^2 \tau_1(\alpha)) \geq g_A(\bar{\tau}_1 \delta_{j-1}) =: \varkappa_1, \quad g_B([1 + \varepsilon]^2 \tau_2(\alpha)) \geq g_B(\bar{\tau}_2 \delta_{k-1}) =: \varkappa_2. \quad (4.31)$$

Moreover, the definition of sets $\Theta'_1(\cdot)$ and $\Theta'_2(\cdot)$ implies that for any $\alpha \in \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}$

$$\Theta_\alpha \subseteq \Theta'_1(\bar{\tau}_1 \delta_j) \cap \Theta'_2(\bar{\tau}_2 \delta_k)$$

and, therefore, for any $\alpha \in \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}$

$$\sup_{\theta \in \Theta_\alpha} \Psi(\chi_\theta) \leq \sup_{\theta \in \Theta'_1(\bar{\tau}_1 \delta_j) \cap \Theta'_2(\bar{\tau}_2 \delta_k)} \Psi(\chi_\theta). \quad (4.32)$$

Probability bound We get from (4.30) and (4.31) for any $\alpha \in \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}$

$$\widehat{U}^{(z, \varepsilon, 0)}(\alpha) \geq U_s^{(\varepsilon)}\left(y, \varkappa, \Theta'_1(\bar{\tau}_1 \delta_j), \Theta'_2(\bar{\tau}_2 \delta_k)\right), \quad y = (1 + \varepsilon)^2 \left[z + R_0(\bar{\tau}_1 \delta_j, \bar{\tau}_2 \delta_k) \right],$$

where we have used that R_0 is increasing/decreasing, in both arguments, function.

It yields together with (4.32) $\forall j, k \geq 0$

$$\begin{aligned} \Psi_{j,k}^{(V)}(z) &:= \sup_{\alpha \in \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}} \left(\sup_{\theta \in \Theta_\alpha} \left\{ \Psi(\chi_\theta) - \widehat{U}^{(z, \varepsilon, 0)}(\alpha) \right\} \right) \\ &\leq \sup_{\theta \in \Theta'_1(\bar{\tau}_1 \delta_j) \cap \Theta'_2(\bar{\tau}_2 \delta_k)} \Psi(\chi_\theta) - U_s^{(\varepsilon)}\left(y, \varkappa, \Theta'_1(\bar{\tau}_1 \delta_j), \Theta'_2(\bar{\tau}_2 \delta_k)\right). \end{aligned}$$

Let us remark that the definition of the sets $\Theta'_A(\cdot)$, $\Theta'_B(\cdot)$ and the functions g_A and g_B as well as their monotonicity imply that

$$\varkappa_1 := g_A(\bar{\tau}_1 \delta_{j-1}) > g_A(\bar{\tau}_1 \delta_j) \geq g_A^*(\bar{\tau}_1 \delta_j) =: \sup_{\theta \in \Theta'_1(\bar{\tau}_1 \delta_j)} A(\theta); \quad (4.33)$$

$$\varkappa_2 := g_B(\bar{\tau}_2 \delta_{k-1}) > g_B(\bar{\tau}_2 \delta_k) \geq g_B^*(\bar{\tau}_2 \delta_k) =: \sup_{\theta \in \Theta'_2(\bar{\tau}_2 \delta_k)} B(\theta), \quad (4.34)$$

and, therefore, Lemma 3 is applicable with $\Theta_1 = \Theta'_1(\bar{\tau}_1 \delta_j)$ and $\Theta_2 = \Theta'_2(\bar{\tau}_2 \delta_k)$.

Thus, applying the first assertion of Lemma 3, we obtain $\forall j, k \geq 0$ and $\forall z \geq 1$

$$\mathbb{P} \left\{ \Psi_{j,k}^{(V)}(z) \geq 0 \right\} \leq 2c \exp \left\{ -z - R_0(\bar{\tau}_1 \delta_j, \bar{\tau}_2 \delta_k) \right\}.$$

Noting that the union of $\left\{ \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}, j = \overline{0, J}, k = \overline{0, K} \right\}$ covers \mathfrak{A} , summing up the right hand side in the last inequality over j, k we come to the first statement of the proposition.

Moment's bound We get from (4.30) and (4.31) for any $\alpha \in \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}$

$$\widehat{U}^{(z, \varepsilon, q)}(\alpha) \geq U_s^{(\varepsilon)}\left(y, \varkappa, \Theta'_1(\bar{\tau}_1 \delta_j), \Theta'_2(\bar{\tau}_2 \delta_k)\right),$$

where $y = (1 + \varepsilon)^2 \left[z + \tau_\varepsilon e_s \left(\varkappa, \Theta'_1(\bar{\tau}_1 \delta_j), \Theta'_2(\bar{\tau}_2 \delta_k) \right) + R_q(\bar{\tau}_1 \delta_j, \bar{\tau}_2 \delta_k) \right]$, $\tau_\varepsilon = 2(2 + \varepsilon)\varepsilon^{-1}$.

It yields together with (4.32) $\forall j, k \geq 0$

$$\begin{aligned}\Phi_{j,k}^{(U)}(z) &:= \left\{ \sup_{\alpha \in \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}} \left(\sup_{\theta \in \Theta_\alpha} \left\{ \Psi(\chi_\theta) - \widehat{U}^{(z, \epsilon, q)}(\alpha) \right\} \right) \right\}_+ \\ &\leq \left\{ \sup_{\theta \in \Theta'_1(\bar{\tau}_1 \delta_j) \cap \Theta'_2(\bar{\tau}_2 \delta_k)} \left[\Psi(\chi_\theta) - U_s^{(\epsilon)}(y, \varkappa, \Theta'_1(\bar{\tau}_1 \delta_j), \Theta'_2(\bar{\tau}_2 \delta_k)) \right] \right\}_+.\end{aligned}$$

Taking into account (4.33), (4.34) and applying the second assertion of Lemma 3, we have, analogously to (4.24), $\forall j, k \geq 0$ and $\forall z \geq 1$

$$\begin{aligned}& \mathbb{E} \left(\Phi_{j,k}^{(U)}(z) \right)^q \\ & \leq 2c\Gamma(q+1) \left[\frac{U_s^{(\epsilon)}(y, \varkappa, \Theta'_1(\bar{\tau}_1 \delta_j), \Theta'_2(\bar{\tau}_2 \delta_k))}{z + \tau_\varepsilon e_{\bar{s}}(\varkappa, \Theta'_1(\bar{\tau}_1 \delta_j), \Theta'_2(\bar{\tau}_2 \delta_k)) + R_q(\bar{\tau}_1 \delta_j, \bar{\tau}_2 \delta_k)} \right]^q \exp \{ -z - R_q(\bar{\tau}_1 \delta_j, \bar{\tau}_2 \delta_k) \}.\end{aligned}$$

Putting for brevity $e_{j,k} = e_{\bar{s}}(\varkappa, \Theta'_1(\bar{\tau}_1 \delta_j), \Theta'_2(\bar{\tau}_2 \delta_k))$ we note that for any $j, k \geq 0$ and $z \geq 1$

$$U_s^{(\epsilon)}(y, \varkappa, \Theta'_1(\bar{\tau}_1 \delta_j), \Theta'_2(\bar{\tau}_2 \delta_k)) \leq 2[g_A(\bar{\tau}_1 \delta_{j-1}) \vee g_B(\bar{\tau}_2 \delta_{k-1})] [4\varepsilon^{-2}e_{j,k} + y]$$

Repeating the computation done after (4.24) we come to the second assertion of the proposition. ■

4.4. Proof of Lemma 1

1⁰. We start the proof with the following simple fact. Let $\widetilde{\mathfrak{T}}$ be an arbitrary subset of \mathfrak{T} . Then

$$\sup_{t \in \widetilde{\mathfrak{T}}} \zeta(t, \cdot) \text{ is } \mathfrak{B} - \text{measurable.} \quad (4.35)$$

Indeed, since \mathfrak{T} is totally bounded $\widetilde{\mathfrak{T}}$ is totally bounded as well. Denote by $\widehat{\mathfrak{T}}$ the union of 2^{-l} -nets, $l \geq 0$, in $\widetilde{\mathfrak{T}}$. Let $\Omega_0 = \{\omega \in \Omega : \zeta(\cdot, \omega) \text{ is continuous}\}$ and let $\overline{\Omega}_0$ be the complementary to Ω_0 . We have for any $x \in \mathbb{R}$

$$\left\{ \omega \in \Omega : \sup_{t \in \widehat{\mathfrak{T}}} \zeta(t, \omega) \leq x \right\} \cap \Omega_0 = \left\{ \omega \in \Omega : \sup_{t \in \widehat{\mathfrak{T}}} \zeta(t, \omega) \leq x \right\} \cap \Omega_0 \in \mathfrak{B}$$

since $\widehat{\mathfrak{T}}$ is countable dense subset of $\widetilde{\mathfrak{T}}$. It remain to note that $\left\{ \omega \in \Omega : \sup_{t \in \widetilde{\mathfrak{T}}} \zeta(t, \omega) \leq x \right\} \cap \overline{\Omega}_0 \in \mathfrak{B}$ since $\mathbf{P}(\overline{\Omega}_0) = 0$ and the considered probability space is complete.

2⁰. Set $\mathfrak{Z}(n, k) = \{\mathfrak{z} \in \mathfrak{Z} : g(\mathfrak{z}) \in [k/n, (k+1)/n]\}$, $n \in \mathbb{N}^*$, $k \in \mathbb{Z}$, and let $\mathbf{K}(n) \subseteq \mathbb{Z}$, $n \in \mathbb{N}^*$, be defined from the relation if $k \in \mathbf{K}(n) \Leftrightarrow \mathfrak{Z}(n, k) \neq \emptyset$. Put also $\mathfrak{T}(k, n) = \cup_{\mathfrak{z} \in \mathfrak{Z}(n, k)} \mathfrak{T}_{\mathfrak{z}}$ and define

$$\begin{aligned}\xi_{k,n}(\omega) &= \sup_{t \in \mathfrak{T}(k,n)} \zeta(t, \omega) - (k+1)/n, & \xi_n(\omega) &= \sup_{k \in \mathbf{K}(n)} \xi_{k,n}(\omega). \\ \eta_{k,n}(\omega) &= \sup_{\mathfrak{z} \in \mathfrak{Z}(n, k)} \left[\sup_{t \in \mathfrak{T}_{\mathfrak{z}}} \zeta(t, \omega) - g(\mathfrak{z}) \right], & \eta(\omega) &= \sup_{\mathfrak{z} \in \mathfrak{Z}} \left[\sup_{t \in \mathfrak{T}_{\mathfrak{z}}} \zeta(t, \omega) - g(\mathfrak{z}) \right].\end{aligned}$$

Some remarks are in order. First, the definition of $\mathfrak{Z}(k, n)$ implies that for any $k \in \mathbf{K}(n)$, $n \in \mathbb{N}^*$

$$\xi_{k,n}(\cdot) \leq \eta_{k,n}(\cdot) \leq \xi_{k,n}(\cdot) + n^{-1}. \quad (4.36)$$

Next, taking into account that $\mathfrak{Z} = \cup_{k \in \mathbf{K}(n)} \mathfrak{Z}(k, n)$ for any $n \in \mathbb{N}^*$ we have

$$\eta(\cdot) = \sup_{k \in \mathbf{K}(n)} \eta_{k,n}(\cdot), \quad \forall n \in \mathbb{N}^*. \quad (4.37)$$

We obtain from (4.36) and (4.37) that for any $n \in \mathbb{N}^*$

$$0 \leq \eta(\cdot) - \xi_n(\cdot) \leq \sup_{k \in \mathbf{K}(n)} [\eta_{k,n}(\cdot) - \xi_{k,n}(\cdot)] \leq n^{-1},$$

and, therefore, $\eta(\cdot) = \lim_{n \rightarrow \infty} \xi_n(\cdot)$. It remains to note that $\xi_{k,n}(\cdot)$ are \mathfrak{B} -measurable for any $k \in \mathbf{K}(n)$, $n \in \mathbb{N}^*$ in view of (4.35), that implies obviously that $\xi_n(\cdot)$ is \mathfrak{B} -measurable for any $n \in \mathbb{N}^*$. Thus, $\eta(\cdot)$ is \mathfrak{B} -measurable as a pointwise limit of \mathfrak{B} -measurable functions. ■

5. Proof of Theorem 1

Below c_1, c_2, \dots , denote the constants completely determined by d, p, μ and γ . We break the proof on several steps.

¹⁰. Let \mathbb{B}_q , $1 < q < \infty$, denote the set of functions vanishing outside \mathbb{K}_μ and those \mathbb{L}_q -norm is less or equal to 1. Later on the integration is always understood as the integration over \mathbb{R}^d .

Consider the set of functions

$$\Theta_h = \left\{ \theta : \mathbb{R}^d \rightarrow \mathbb{R} : \theta(x) = h^{-d} \int K_h(t-x) \ell(t) dt, \quad \ell \in \mathbb{B}_{\frac{p}{p-1}} \right\}, \quad h \in \mathcal{H}.$$

Put also $\Theta = \cup_{h \in \mathcal{H}} \Theta_h$ and for any $\theta \in \Theta$ introduce the gaussian process

$$\chi_\theta = \int_{\mathbb{R}^d} \theta(x) b(dx). \quad (5.1)$$

In view of Young inequality Folland (1999), Theorems 6.18 and 6.36, for any $\theta \in \Theta$

$$\|\theta\|_2 := \left(\int_{\mathbb{R}^d} |\theta(x)|^2 dx \right)^{\frac{1}{2}} \leq h^{\frac{d(2-p)}{2p}} \|K\|_{\frac{2p}{p+2}} \leq h^{\frac{d(2-p)}{2p}} \leq (h^{(\min)})^{\frac{d(2-p)}{2p}} < \infty, \quad (5.2)$$

since $2 \leq p < \infty$. Here we have also used that $\|K\|_q \leq \|K\|_\infty \leq 1$ in view of assumption imposed on the function K .

Thus, the stochastic integral in (5.1) is well-defined and χ_θ is zero-mean gaussian random function on Θ such that

$$V(\theta) = \|\theta\|_2, \quad \rho(\theta_1, \theta_2) = \|\theta_1 - \theta_2\|_2. \quad (5.3)$$

We equip Θ with the semi-metric ρ and in the next paragraph we compute the entropy of several subsets of Θ . This computation allows us, in particular, to assert that Dudley integral is finite on Θ . It yields Lifshits (1995) that χ_\bullet is P-a.s uniformly continuous on (Θ, ρ) , therefore Assumption 2 holds. Moreover, we show that Assumption 3 is fulfilled.

We conclude that Proposition 3 are applicable to χ_θ , on Θ , with $\Theta_\alpha = \Theta_h$, $\mathfrak{A} = \mathcal{H}$, $\mathfrak{a} = \sqrt{2}\rho$, $A = \sqrt{2}V$. It remains to note that in view of duality arguments for any $h \in \mathcal{H}$

$$\|\xi_h\|_p = \sup_{\ell \in \mathbb{B}_{\frac{p}{p-1}}} \int \ell(t) \xi_h(t) dt,$$

and, therefore, for any $h \in \mathcal{H}$

$$\|\xi_h\|_p = \sup_{\theta \in \Theta_h} \chi_\theta. \quad (5.4)$$

2⁰. In order to apply Proposition 3 we need to compute several quantities. First, we have to choose the function τ_1 (since $B, \mathfrak{b} \equiv 0$, hence $\tau_2 \equiv 0$). Set $\tau_1(h) = h^{-1}$ and note that for any $u > 0$

$$\Theta'_1(u) := \bigcup_{h: \tau_1(h) \leq u} \Theta_h = \bigcup_{h \geq u^{-1}} \Theta_h,$$

We note that computations given in (5.2) yield

$$g_A^*(u) := \sup_{\theta \in \Theta'_1(u)} A(\theta) \leq \sqrt{2} u^{\frac{d(p-2)}{2p}} =: g_A(u), \quad u \geq \left(h^{(\max)}\right)^{-1}.$$

3⁰. Let $\mathfrak{E}^{(u)}(\delta)$, $\delta > 0$, be the entropy of $\Theta'_1(u)$ computed with respect to semi-metric $\mathfrak{a} = \sqrt{2}\rho$, where, remind $\rho(\cdot) = \|\cdot\|_2$. The following assertion is true: there exist c_0 completely determined by γ, d, p and μ such that for any $\forall \bar{\gamma} \in (d/2 - d/p, \gamma]$ and for any $u \geq \left(h^{(\max)}\right)^{-1}$

$$\mathfrak{E}^{(u)}(\delta) \leq c_1 u^d \delta^{-d/\bar{\gamma}}, \quad \forall \delta > 0. \quad (5.5)$$

where $c_1 = c_0 3^{4d} d^2$.

The proof of (5.5) is obtained by routine computations and it is postponed to the step 8⁰.

4⁰. Choosing $u = \left(h^{(\min)}\right)^{-1}$ (that yields $\Theta'_1(u) = \Theta$) and $\bar{\gamma} = \gamma$ we get for any $\delta > 0$

$$\mathfrak{E}_{\Theta, \mathfrak{a}}(\delta) \leq c_1 \left(h^{(\min)}\right)^{-d} \left(\frac{1}{\delta}\right)^{d/\gamma}.$$

In view the condition $\gamma > d/2$, Dudley integral is finite on Θ and, therefore, χ_\bullet is P-a.s uniformly continuous on (Θ, ρ) . It complete the verification of Assumption 2.

The last inequality shows also that there exist $\tau > 0$ such that for any $s \in \mathbb{S}$, satisfying $\limsup_{a \rightarrow \infty} a^\tau s^{-1}(a) < \infty$, Assumption 3 is fulfilled.

5⁰. Let us now choose the function \mathfrak{s} . Set $\delta(u) = u^{d/2-d/p}$ and for any $u \geq 1$ let $m(u) \in \mathbb{N}$ be such that $2^{m(u)} \leq \delta(u) < 2^{m(u)+1}$. Define

$$\mathfrak{s}(u, \delta) = (3/4\pi^2) \left(1 + \left[\log_2(2^{-m(u)}\delta)\right]^2\right)^{-1}.$$

We remark that

$$\mathfrak{s}(u, 2^{k/2}) = (3/4\pi^2) \left[1 + ((k/2) - m(u))^2\right]^{-1}$$

and, therefore, $\mathfrak{s}(u, \cdot) \in \mathbb{S}$ for any $u \geq 1$. Moreover, we note that if $p = 2$ then \mathfrak{s} does not depend on u and it is given by

$$\mathfrak{s}(\delta) = (3/4\pi^2) \left(1 + [\log_2 \delta]^2\right)^{-1}.$$

Obviously the factor $3/4\pi^2$ can be replaced here by $6/\pi^2$.

Now, let us compute the quantity λ_1 related to the function \mathfrak{s} . Remind that

$$\lambda_1 := \sup_{t \in [1, \sqrt{2}]} \sup_{x > \tau_1} \sup_{\delta > 0} \frac{\mathfrak{s}(xt, \delta)}{\mathfrak{s}(x, \delta)},$$

where $\tau_1 = \left(h^{(\max)}\right)^{-1} \geq 1$. It is evident that

$$\begin{aligned} \lambda_1 &= \sup_{m \geq 0} \left\{ 1 \vee \sup_{\delta > 0} \left[\frac{1 + \left(\log_2(2^{-m}\delta)\right)^2}{1 + \left(\log_2(2^{-m-1}\delta)\right)^2} \right] \right\} = 1 \vee \sup_{x > 0} \left[\frac{1 + \left(\log_2(x)\right)^2}{1 + \left(\log_2(x/2)\right)^2} \right] \\ &\leq 1 + \sup_{x > 0} \left| \frac{2 \log_2(x) - 1}{1 + \left(\log_2(x) - 1\right)^2} \right| = 1 + \sup_{y \in \mathbb{R}} \left| \frac{2y + 1}{1 + y^2} \right| < 3. \end{aligned}$$

6⁰. Define

$$\bar{\gamma}(\delta) = \begin{cases} d/2 - d/(2p), & 0 < \delta < \delta(u); \\ \gamma, & \delta \geq \delta(u), \end{cases}$$

and note that $(d/2 - d/p) < \gamma(\delta) \leq \gamma$ for any $\delta > 0$.

Putting $c_2 = c_1 4(144)^4$ we get for any $\delta > 0$ in view of (5.5)

$$\begin{aligned} \mathfrak{E}^{(u)} \left(\lambda_1^{-1} [g_A(u) \mathfrak{s}(u, \delta)] / 48\delta \right) &\leq c_2 u^d \left[u^{\frac{d(p-2)}{2p}} \mathfrak{s}(u, \delta) \right]^{-d/\bar{\gamma}(\delta)} \delta^{d/\bar{\gamma}(\delta)} \\ &= c_2 u^{\frac{2d}{p} - \left(\frac{d(p-2)}{2p}\right) \left(\frac{d}{\bar{\gamma}(\delta)} - 2\right)} \left[\mathfrak{s}(u, \delta) \right]^{-d/\bar{\gamma}(\delta)} \delta^{d/\bar{\gamma}(\delta)} \\ &= c_2 u^{\frac{2d}{p}} [\delta(u)]^{2 - \frac{d}{\bar{\gamma}(\delta)}} \left[\mathfrak{s}(u, \delta) \right]^{-4} \delta^{d/\bar{\gamma}(\delta)}. \end{aligned} \quad (5.6)$$

To get the last inequality we have taken into account that $\mathfrak{s}(u, \delta) < 1$ for any $u \geq 1, \delta > 0$ and that $d/\bar{\gamma}(\delta) \leq 4$.

We obtain from (5.6) for any $\delta > 0$, putting $c_3 = (4\pi^2/3)^4 c_2$,

$$\begin{aligned} \delta^{-2} \mathfrak{E}^{(u)} \left(\lambda_1^{-1} [g_A(u) \mathfrak{s}(u, \delta)] / 48\delta \right) &\leq c_2 u^{\frac{2d}{p}} \left(\frac{\delta}{\delta(u)} \right)^{d/\bar{\gamma}(\delta) - 2} \left[\mathfrak{s}(u, \delta) \right]^{-4} \\ &\leq c_2 u^{\frac{2d}{p}} \left[\left(\delta / \delta(u) \right)^{2/(p-1)} 1_{(0, \delta(u))}(\delta) + (\delta(u)/\delta)^{2-d/\gamma} 1_{[\delta(u), \infty)}(\delta) \right] \left[\mathfrak{s}(u, \delta) \right]^{-4} \\ &\leq c_2 u^{\frac{2d}{p}} \left[\left(2^{-m(u)} \delta \right)^{2/(p-1)} 1_{(1, \delta(u))}(\delta) + 4 \left(2^{m(u)} / \delta \right)^{2-d/\gamma} 1_{[\delta(u), \infty)}(\delta) \right] \left[\mathfrak{s}(u, \delta) \right]^{-4} \\ &= 5c_3 u^{\frac{2d}{p}} 2^{-\alpha |\log_2(2^{-m(u)} \delta)|} \left(1 + \left[\log_2(2^{-m(u)} \delta) \right]^2 \right)^4, \end{aligned} \quad (5.7)$$

where $\alpha = \min \{2/(p-1), 2-d/\gamma\}$. We obtain from (5.7)

$$\mathcal{E}'(u) := \sup_{\delta > 0} \left[\delta^{-2} \mathfrak{E}^{(u)} \left(\lambda_1^{-1} [g_A(u) \mathfrak{s}(u, \delta)] / 48\delta \right) \right] \leq c_4 u^{\frac{2d}{p}}, \quad (5.8)$$

where $c_4 = 5c_3 \sup_{x \geq 0} \left[2^{-\alpha x} (1+x^2)^4 \right]$.

7⁰. Remind that $\tau_1(h) = h^{-1}$ and, in particular, $\bar{\tau}_1 = [h^{(\min)}]^{-1}$. Choosing $\varepsilon = \sqrt{2} - 1$ we get from (5.8) putting $c_5 = c_4 2^{d/p}$

$$\hat{\mathcal{E}}^{(\sqrt{2}-1)}(h) := \mathcal{E}'(\sqrt{2}\tau_1(h)) \leq c_5 h^{-\frac{2d}{p}}.$$

Choose also $R_r(t) = R(t) = t^{\frac{2d}{p}}$ (independent of r) that yields

$$\hat{R}^{(\sqrt{2}-1)}(h) := R(\sqrt{2}\tau_1(h)) = 2^{d/p} h^{-\frac{2d}{p}}.$$

Choosing finally $z = 0$ and putting $\hat{U}(h) = \hat{U}^{(z, \varepsilon, r)}(h)$, $\varepsilon = \sqrt{2} - 1$, $z = 1$, we obtain that

$$\hat{U}(h) \leq \sqrt{2} g_A(2\tau_1(h)) \sqrt{32c_5 h^{-\frac{2d}{p}} + 2^{d/p} h^{-\frac{2d}{p}} + 1} \leq c_6 h^{\frac{d(2-p)}{2p} - \frac{d}{p}} = c_6 h^{-\frac{d}{2}}.$$

Let us compute now the quantities $\mathcal{R}^{(\varepsilon, 0)}$, $\mathcal{R}^{(\varepsilon, q)}$ defined in (2.10) with $\varepsilon = \sqrt{2} - 1$.

Noting that $\underline{\tau}_1 = (h^{(\max)})^{-1}$ and $2^{J/2} \bar{\tau}_1 \leq 2^{-1/2} \underline{\tau}_1$ we get

$$\begin{aligned} \mathcal{R}^{(\varepsilon, 0)} &:= \sum_{j=0}^J \exp \left\{ - \left(\bar{\tau}_1 2^{-(j/2)} \right)^{\frac{2d}{p}} \right\} \leq c_8 \exp \left\{ - 2^{-3/2} \left(h^{(\max)} \right)^{-2d/p} \right\}; \\ \mathcal{R}^{(\varepsilon, q)} &:= \sum_{j=0}^J \left[g_A(\bar{\tau}_1 2^{-j/2}) \right]^q \exp \left\{ - \left(\bar{\tau}_1 2^{-(j/2)} \right)^{\frac{2d}{p}} \right\} \\ &\leq c_9 \left(h^{(\max)} \right)^{\frac{qd(2-p)}{2p}} \exp \left\{ - 2^{-3/2} \left(h^{(\max)} \right)^{-2d/p} \right\}. \end{aligned}$$

The assertions of the theorem follow now from Proposition 3.

8⁰. It remains to prove (5.5). The proof is based on the following inclusion: for any $\bar{\gamma} \in (0, \gamma]$

$$\Theta'_1(u) \subset \mathbb{H}_{\frac{p}{p-1}}^* \left(\bar{\gamma}, 3^d \sqrt{d} u^{\bar{\gamma}} \right), \quad \forall u > 0, \quad (5.9)$$

where $\mathbb{H}_q^*(\cdot, \cdot) \subset \mathbb{H}_q(\cdot, \cdot)$ consists of functions vanishing outside of $\mathbb{K}_{2\mu}$.

Let $\mathfrak{E}^*(\cdot)$ be the entropy of $\mathbb{H}_{\frac{p}{p-1}}^*(\bar{\gamma}, L)$, $L > 0$, measured in $\|\cdot\|_2$. It is well-known Edmunds and Triebel (1996), that for any $p > 1$ there exist c_0 completely determined by γ, d, p and μ such that for any $(d/2 - d/p) < \bar{\gamma} \leq \gamma$ and for any $L > 0$

$$\mathfrak{E}^*(\delta) \leq c_0 \left(L \delta^{-1} \right)^{d/\bar{\gamma}}, \quad \forall \delta > 0.$$

Since $\bar{\gamma}$ takes only two values $d/2 - d/2p$ and γ , (5.5) follows immediately from (5.9). Thus, we shall prove (5.9).

Fix $\theta \in \Theta'_1(u)$. By its definition there exists $\ell \in \mathbb{B}_{\frac{p}{p-1}}$ and $h \geq u^{-1}$ such that $\theta = K_h * \ell$, where “ $*$ ” stands convolution operator on \mathbb{R}^d . First, we note that all functions belonging to Θ vanish outside the cube $\mathbb{K}_{2\mu}$ in view of assumption imposed on function K .

Next, for any $m = (m_1, \dots, m_d) \in \mathbb{N}^d$ put $|m| = m_1 + \dots + m_d$, and set $\gamma = l + \alpha$ and $\bar{\gamma} = \bar{l} + \bar{\alpha}$, where $l, \bar{l} \in \mathbb{N}$ and $0 < \alpha, \bar{\alpha} \leq 1$.

Then, since $K \in \mathbb{H}_\infty(\gamma, 1)$ we have for any $m \in \mathbb{N}^d$ such that $|m| \leq l$

$$\sup_{x \in \mathbb{R}^d} \left| \frac{\partial^{|m|} K_h(x)}{\partial^{m_1} x_1 \dots \partial^{m_d} x_d} \right| \leq h^{-|m|} \leq (h^{(\min)})^{-\gamma} < \infty.$$

Above remarks allow us to assert that all partial derivatives $\theta^{(m)}$ exist whenever $|m| \leq l$ and they are given by

$$\theta^{(m)}(x) = \int (K_h)^{(m)}(t - x) \ell(t) dt, \quad \forall x \in \mathbb{R}^d,$$

where for any function g the notation $g^{(m)}$ or (if it is more convenient) $(g)^{(m)}$ is used for its partial derivative.

We obtain in view of Young inequality for any $\Delta \in \mathbb{R}$ and any $m \in \mathbb{N}^d$ satisfying $|m| = \bar{l}$

$$\begin{aligned} \left\| \theta^{(m)}(\cdot + \Delta) - \theta^{(m)}(\cdot) \right\|_{\frac{p}{p-1}} &\leq \left\| (K_h)^{(m)}(\cdot + \Delta) - (K_h)^{(m)}(\cdot) \right\|_1 \\ &\leq h^{-\bar{l}} \left\| K^{(m)}(\cdot + [\Delta/h]) - K^{(m)}(\cdot) \right\|_1. \end{aligned}$$

Here we have used that $\ell \in \mathbb{B}_{\frac{p}{p-1}}$.

We remark that if $h \leq |\Delta|$ then for any $u \in \mathbb{R}^d$ either $K^{(m)}(u + [\Delta/h]) = 0$ or $K^{(m)}(u) = 0$ in view of the assumption imposed on the support of K . Thus, if $h \leq |\Delta|$

$$\left\| K^{(m)}(\cdot + [\Delta/h]) - K^{(m)}(\cdot) \right\|_1 \leq \|K^{(m)}\|_1 \leq \|K^{(m)}\|_1 (\Delta/h)^{\bar{\alpha}} \leq (\Delta/h)^{\bar{\alpha}},$$

since $h \leq |\Delta|$ and $\|K^{(m)}\|_1 \leq \|K^{(m)}\|_\infty \leq 1$ in view of assumption imposed on the function K .

If $h > |\Delta|$ then in view of the assumption imposed on the support of K we have

$$\begin{aligned} \left\| K^{(m)}(\cdot + [\Delta/h]) - K^{(m)}(\cdot) \right\|_1 &= \int_{[-\frac{3}{2}, \frac{3}{2}]^d} \left| K^{(m)}(u + [\Delta/h]) - K^{(m)}(u) \right| du \\ &\leq 3^d \left| \sum_{i=1}^d (\Delta/h)^2 \right|^{\bar{\alpha}/2} \leq 3^d \sqrt{d} (|\Delta|/h)^{\bar{\alpha}}. \end{aligned}$$

Since $h \geq u^{-1}$ we conclude finally that $\forall \Delta \in \mathbb{R}$

$$\left\| \theta^{(m)}(\cdot + \Delta) - \theta^{(m)}(\cdot) \right\|_{\frac{p}{p-1}} \leq 3^d \sqrt{d} h^{-\bar{\gamma}} |\Delta|^{\bar{\alpha}} \leq 3^d \sqrt{d} u^{\bar{\gamma}} |\Delta|^{\bar{\alpha}}. \quad (5.10)$$

It means that $\theta \in \mathbb{H}_{\frac{p}{p-1}}(\bar{\gamma}, 3^d \sqrt{d} u^{\bar{\gamma}})$. As it was mentioned above all function belonging to Θ vanish outside the cube $\mathbb{K}_{2\mu}$ that allows us to conclude that $\theta \in \mathbb{H}_{\frac{p}{p-1}}^*(\bar{\gamma}, 3^d \sqrt{d} u^{\bar{\gamma}})$ and, therefore, (5.9) is proved. ■

References

- ALEXANDER, K. S. (1984). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* **12**, 1041–1067.
- BARRON, A., BIRGÉ, L. and MASSART, P. (1999). Risk bounds for model selection via penalization. *Probab. Theory Related Fields* **113**, 301–413.
- BOBKOV, S.G. (1988). Upper functions and oscillating Gaussian processes. *J. Soviet Math.* **43**, **6**, 2745–2751.
- BOUSQUET, O. (2002). A Bennett concentration inequality and its application to suprema of empirical processes. *C. R. Math. Acad. Sci. Paris* **334**, 495–500.
- CAVALIER, L. and GOLUBEV, YU. (2006). Risk hull method and regularization by projections of ill-posed inverse problems. *Ann. Statist.* **34**, 1653–1677.
- COIFMAN, R.R. and WEISS, G. (1971). *Analyse harmonique non-commutative sur certaines espaces homogènes*. Lectures Notes in Math. **242**, Springer-Verlag, Berlin.
- COULHON, T. , KERKYACHARIAN, G. and PETRUSHEV, P. (2011). Heat kernel generated frames in the setting of Dirichlet spaces. *Manuscript*.
- DONY, J. , EINMAHL, U. and MASON, D. (2006). Uniform in bandwidth consistency of local polynomial regression function estimators. *Australian Journal of Stat.* **35**, 105–120.
- DONY, J. and EINMAHL, U. (2009). Uniform in bandwidth consistency of kernel regression estimators at a fixed point. *IMS Collection* **5**, (Highdimensional probability. The Lumini volume), 308–325.
- EDMUNDS, D.E. and TRIEBEL, H. (1996). Function Spaces, Entropy numbers and Differential Operators. *Cambridge University Press*
- EGISHYANTS, S.A. and OSTROVSKII, E. I. (1996). Local and global upper function for random fields. *Theory Probab. App.* **41**, **4**, 657–665.
- EINMAHL, U. and MASON, D. M. (2000). An empirical process approach to the uniform consistency of kernel-type function estimators. *J. Theoret. Probab.* **13**, 1–37.
- EINMAHL, U. and MASON, D. M. (2005). Uniform in bandwidth consistency of kernel-type function estimators. *Annals Stat.* **33**, **3**, 1380–1403.
- FOLLAND, G. B. (1999). *Real Analysis. Modern Techniques and Their Applications*. Second edition. John Wiley & Sons, Inc., New York.
- GINÉ, E. and KOLTCHINSKII, V. (2006). Concentration inequalities and asymptotic results for ratio type empirical processes. *Ann. Probab.* **34**, 1143–1216.
- GINÉ, E. and ,GUILLOU A. (2002). Rate of strong uniform consistency for multivariate kernel density estimation. *Ann. Inst. H.Poincaré, Proba.Statist.* **38**, 907–921.
- GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12**, 928–989.
- GOLDENSHLUGER, A. and LEPSKI, O. (2008). Universal pointwise selection rule in multivariate function estimation. *Bernoulli* **14**, 1150–1190.
- GOLDENSHLUGER, A. and LEPSKI, O. (2009). Structural adaptation via L_p -norm oracle inequalities. *Probability Theory and Related Fields* **143**, 41–71.
- GOLDENSHLUGER, A. and LEPSKI, O. (2011). Uniform bounds for norms of sums of independent random functions. *Ann. Probab.* **39**, **6**, 2318–2384.
- GOLUBEV, Y. and SPOKOINY, V. (2009). Exponential bounds for minimum contrast estimators. *Electron. J. Stat.* **3**, 712–746.
- KALINAUSKAITĖ, N. (1966). Upper and lower functions for sums of independent random variables with limiting stable distribution. *Litovsk. Mat.Sb.* **6**, 249–256.

- KERKYACHARIAN, G. , LEPSKI, O.V. and PICARD, D. (2001). Non linear estimation in anisotropic multi-index denoising. *Probab. Theory and Relat. Fields* **121**, 137–170.
- KLUCHNIKOFF, N. (2005). On adaptive estimation of anysotropic functions. *Ph.d Thesis. Aix-Marseille 1.* .
- LEPSKI, O. V. (1991). Asymptotically minimax adaptive estimation. I. Upper bounds. Optimally adaptive estimates. *Theory Probab. Appl.* **36**, 682–697.
- LEPSKI, O.V. , MAMMEN, E. and SPOKOINY V.G. (1997). Ideal spatial adaptation to inhomogeneous smoothness: an approach based on kernel estimates with variable bandwidth selection. *Ann. Statist.* **25**, 929–947.
- LEPSKI, O. V. and SPOKOINY, V. G. (1997). Optimal pointwise adaptive methods in nonparametric estimation. *Annals Stat.* **25**, 2512–2546.
- LIFSHTS, M.A. (1995). *Gaussian Random Functions*. Kluwer, Dordrecht.
- MASSART, P. (2000). About the constants in Talagrand’s concentration inequalities for empirical processes. *Ann. Probab.* **28**, 863–884.
- QUALLS, C. and WATANABE, H. (1972). Asymptotic prpperties of Gaussian processes. *Annals Math. Stat.* **43**, 580–596.
- SHIRAYEV, A.N., VALKEILA, E. and VOSTRIKOVA, L. (2002). On lower and upper function for square integrable martingales. *in Proc.Steklov Inst. Math.* **2 (237)**, 281–292.
- SPOKOINY, V. (1996). Adaptive hypothesis testing using wavelets. *Ann. Stat.* **24**, 2477–2498.
- STONE, C.J. (1985). Additive regression and other nonparametric models. *Ann. Stat.* **13, 2**, 689–705.
- TALAGRAND, M. (1994). Sharper bounds for Gaussian and empirical processes. *Ann. Probab.* **22**, 28–76.
- TALAGRAND, M. (2005). *The Generic Chaining*. Springer, New York.
- TSYBAKOV, A.B. (1998). Pointwise and sup-norm sharp adaptive estimation of functions on the Sobolev classes. *Ann. Stat.* **26**, 2420–2469.
- VAN DE GEER, S. (2000). *Applications of Empirical Process Theory*. Cambridge University Press, Cambridge.
- VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer, New York.